NAVPERS 93400A-3

FUNDAMENTALS OF ELECTRONICS

VOLUME 3

TRANSMITTER CIRCUIT APPLICATIONS

PREFACE

This book is part of an nine-volume set entitled "Fundamentals of Electronics". The nine volumes include:

Volume 1a - NavPers 93400A-1a, Basic Electricity, Direct Current
Volume 1b - NavPers 93400A-1b, Basic Electricity, Alternating Curren
Volume 2 - NavPers 93400A-2, Power Supplies and Amplifiers
Volume 3 - NavPers 93400A-3, Transmitter Circuit Applications
Volume 4 - NavPers 93400-4, Receiver Circuit Applications
Volume 5 - NavPers 93400-5, Oscilloscope Circuit Applications
Volume 6 - NavPers 93400-6, Microwave Circuit Applications
Volume 7 - NavPers 93400-7, Electromagnetic Circuits and Devices
Volume 8 - NavPers 93400-8, Tables and Master Index

If you are becoming acquainted with electricity or electronics for the first time, study volumes one through seven in their numerical sequence. If you have a background equivalent to the information contained in volumes one and two, you are prepared to study the material contained in any of the remaining volumes. A master index for all volumes is included in volume eight. Volume eight also contains technical and mathematical tables that are useful in the study of the other volumes.

A question (or questions) follows each group of paragraphs. The questions are designed to determine if you understand the immediately preceding information. As you study, write out your answers to each question on a sheet of paper. If you have difficulty in phrasing an answer, restudy the applicable paragraphs. Do not advance to the next block of paragraphs until you are satisfied that you have written a correct answer.

When you have completed study of the text matter and written satisfactory answers to all questions on two facing pages of the book, compare your answers with those at the top of the next even-numbered page. If the answers match, you may continue your study with reasonable assurance that you have understood and can apply the material you have studied. Whenever your answers are incorrect, restudy the applicable material to determine why the book answer is correct and yours is not. If you make an honest effort to follow these instructions, you will have achieved the maximum learning benefits from each study assignment.

Follow the directions of your instructor in answering the review questions included at the end of each chapter.

TABLE OF CONTENTS

CHAPTER 24

OSCILLATORS

A simple radio transmitter was shown previously to consist of three major blocks, a dc power supply, an AF modulator, and an RF section. The RF section is necessary because it is impractical to transmit audio frequencies over long distances. In order to transmit audio (and other types of intelligence) long distances, it is necessary to combine the intelligence with a radio frequency signal called the CARRIER. The method of combining these signals will be presented at a later time. The carrier is generated in the RF section. The RF section can be subdivided into three blocks, the OSCILLATOR, the INTERMEDIATE POWER AMPLIFIER, and the FINAL POWER AMPLIFIER. The purpose of this chapter is to explain the theory and operation of the oscillator block.

24-1. The Oscillator

A question that might arise at this time is, "What is an oscillator?" This question can best be answered by first determining what an oscillator does. The primary function of an oscillator is to generate a given waveform at a constant amplitude and specific frequency and maintain this waveform within certain limits. The oscillators discussed in this chapter will produce a sinusoidal waveform. A generator or alternator will produce a sinusoidal waveform at a given frequency, (such as 60 cps for the ac power line) and it is possible to maintain the frequency within very strict limits. However, generators and alternators are NOT oscillators. It is true that before the advent of the electron tube some alternators were made that could produce a radio frequency (the Alexanderson alternator produced 100 Kc), but even their top limit was still very low as far as RF is concerned. Along with the disadvantage of their frequency limitations they were very expensive, and could be used at only one frequency. The development of the electron tube solved the problem of generating high frequency ac.

Again the question is asked, "What is an oscillator?" By definition an oscillator is, "a nonrotating device for producing alternating current, the output frequency of which is determined by the characteristics of the device." The definition includes the word nonrotating in order to eliminate the possibility of including an alternator as an oscillator. An alternator is a rotating device and would NOT be classified as a oscillator. The basic definition can be enlarge upon in the following manner: "An electron oscillator is a device that generates alternatin current from a direct current source." Ar amplifier may be used for this purpose. Sin electron tubes are amplifiers they can be us as oscillators: Electron tubes are essential converters that change dc electrical energy fro the plate power supply into ac electrical energy

The mere fact that a circuit is capable oscillating does not make it a useful oscillato In order to qualify as an oscillator, a circu must be able to generate sustained oscillatio in a desired and controllable manner. Mo people are familiar with the undesired oscillati in a public address system (loud whistle) th occurs when the microphone is moved too clo to the speaker.

An electron tube oscillator circuit is esse tially an amplifier that is designed to supply i own input through a feedback network. Tv conditions must be met by this feedback netwo if sustained oscillations are to be produce First, the voltage feedback from the output mu be in phase with the excitation voltage on th grid; that is, the feedback must aid the inpu This is called POSITIVE or REGENERATIV feedback. Second, the amount of energy fee back to the grid circuit must be sufficient compensate for the energy losses in the gr circuit.

Feedback may be accomplished by inductiv capacitive, or resistive coupling between tl plate and the grid circuit. Various circuits hav been developed to produce feedback of the prope phase and amount. Each of these circuits h certain characteristics that make its use advar tageous under given circumstances.

Without the electronic oscillator very fe advanced electronic circuit applications wou be possible. Because of this fact the importan of a thorough understanding of oscillator theo and operation is stressed.

The list of types of equipment employing o cillators (in one form or another) is a long on including equipment such as radar, sonar, guid missiles, communications gear, test equipmer home entertainment products, etc.

24-2. Desired Oscillator Characteristics

Virtually every piece of equipment that utilizes an oscillator has two main requirements of the oscillator-AMPLITUDE STABILITY and FREQUENCY STABILITY. The rigidity with which these requirements must be met depends on the accuracy demanded of the equipment.

Amplitude stability refers to the ability of the oscillator to maintain a constant amplitude output waveform. The less the deviation from a predetermined amplitude, the better is the amplitude stability.

Frequency stability refers to the ability of the oscillator to maintain the desired operating frequency. The less the oscillator drifts from the operating frequency, the better is the frequency stability.

24-3. Review of Parallel Resonance

The parallel resonant circuit, sometimes known as a tank circuit due to its ability to store energy, is the heart of most oscillators. A complete description of tank circuit action is provided in Chapter 12. At this point a brief discussion of tank circuit operation as it pertains to oscillators is given.

Figure 24-1 shows the tank circuit during various phases of its cycle. Assume that the capacitor is fully charged to the battery potential, the battery is disconnected; and the capacitor is acting as the source. At this instant all the circuit energy is stored in the electrostatic field of the capacitor. The tank voltage is equal to the capacitor voltage and the tank current is zero. The capacitor starts to discharge through the inductor, producing a magnetic field around it. At this point, tank voltage is decreasing and tank current is increasing. At the instant (t_2) the capacitor is fully discharged and the following conditions exist: tank voltage is zero, tank current is maximum, the magnetic field surrounding the inductor is maximum. All of the circuit energy is now stored in the magnetic field of the inductor. Since the applied EMF (capacitor voltage) is now zero, tank current tends to stop flowing. The magnetic field around the inductor begins to collapse trying to maintain tank current flow in the same direction. This causes the capacitor to start recharging with the opposite polarity. At this time (t_3) tank current is decreasing, and tank voltage is increasing with the opposite polarity. When the magnetic field has completely collapsed (t_4) tank current is zero, and the capacitor has fully charged. Tank voltage is now maximum and the circuit energy is again stored in the capacitor's electrostatic field. The entire sequence would now occur again with tank current flowing in the opposite direction.

Tank energy is alternately stored in the capacitor's electrostatic field and the inductor's

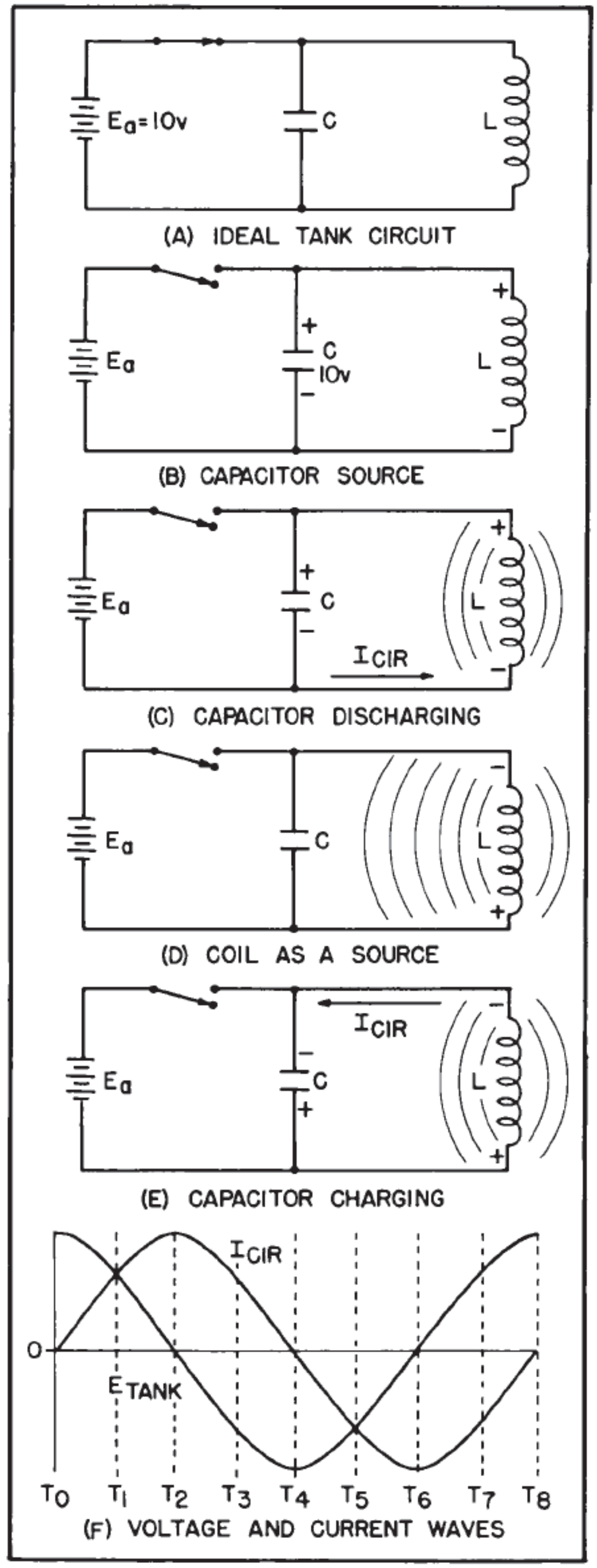


Figure 24-1 - Tank circuit action.

magnetic field. The interchange of energy between the two components is called the FLYWHEEL EFFECT. In an ideal tank circuit (a circuit with no power losses), these oscillations would continue indefinitely. A practical circuit will produce the waveshape shown in Figure 24-2.

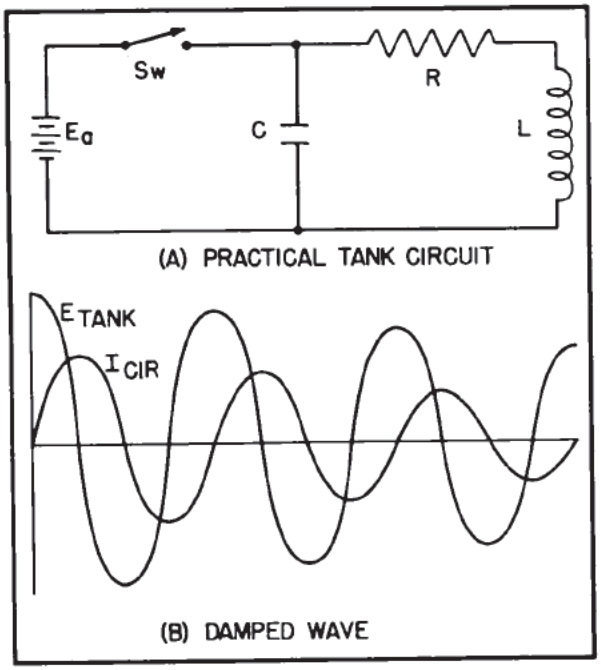


Figure 24-2 - Practical tank circuit.

In this case, the I^2R losses which occur reduce the energy exchanged between L and C producing damped oscillations. The rate of damping is determined by the magnitude of the I^2R losses. The addition of energy with the proper magnitude and at the proper time will allow the practical tank circuit to sustain oscillations.

When the tank circuit is used to develop oscillations, for all practical purposes, the operating frequency of the oscillator is the same as the resonant frequency of the tank circuit and can be found by the formula:

$$fo = \frac{1}{2\pi\sqrt{LC}} \qquad (11\text{-}16)$$

This formula is valid because most of the tank circuits used in RF oscillators have a Q greater than 10.

Q1. Can oscillations be maintained if the energy losses are just a little greater than the energy returned each cycle?

24-4. The Basic Oscillator Circuit

The block diagram of a basic oscillator is shown in Figure 24-3.

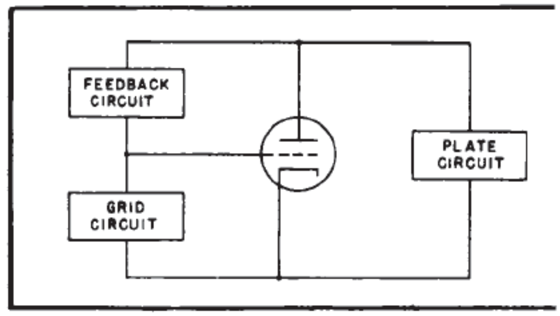


Figure 24-3 - Block diagram of a vacuum tu oscillator circuit.

The basic oscillator can be broken down i three main sections: a frequency determin device, an amplifier, and a feedback circu The frequency determining device is usually LC tank circuit and is located in the grid circu

The electron tube itself is not an oscillat The oscillations actually take place in the tur circuit, a part of which may be composed of interelectrode capacitances of the electron tu and the distributed capacitances and inductanc of the circuit. The electron tube functions p marily as an electrical valve that amplifies, a automatically delivers to the grid circuit proper amount of energy to maintain oscillati The feedback circuit couples energy from plate circuit back to the grid circuit. The fee back circuit may be variable to control amount of energy feedback.

The circuit is essentially a closed loop util zing dc power to maintain the ac oscillation. oscillators, that will be studied in this chapt use this principle of operation; and differ mai in the type and method of feedback used.

SIMPLE ARMSTRONG OSCILLATOR

The simplest type of oscillator is the ARI STRONG OSCILLATOR. Figure 24-4 illustrat a simple Armstrong oscillator with zero bi

L_2; C_1 form the tank circuit which determine the resonant frequency.

V_1 is the oscillator tube which acts as a swi and provides amplification.

L_1 is the feedback coil; sometimes called TICKLER COIL. L_1 serves as the plate lc with feedback being supplied by the inducti coupling between L_1 and L_2.

Switch S_1 is inserted merely as an aid to e planation. Once oscillations have been star in the tank circuit they appear in amplified fo in the plate circuit of the tube. Since the ele tron tube is a voltage amplifying device, voltage of the tank circuit will be the controlli factor in the operation of the circuit. It has be

A1. No.

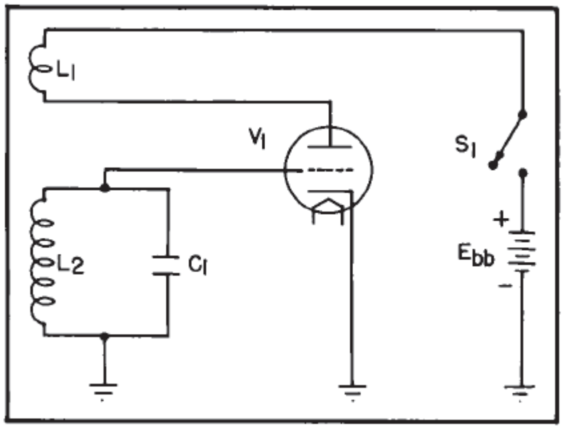


Figure 24-4 - Simple armstrong oscillator.

stated previously that the voltage that appears across the capacitor is the tank voltage. Therefore, a discussion of circuit operation will be concerned primarily with the charge and discharge of the tank capacitor.

Before entering into a detailed analysis of operation there are two important concepts which must be reestablished.

The first concept is that of reflected impedance. Figure 24-5A shows a schematic diagram of a transformer with an unloaded secondary. Two meters are connected to the transformer: an ammeter in the primary to measure the primary current, and a voltmeter across the secondary to measure the voltage across the secondary winding. A switch is provided in the secondary to permit connection of the load. The load in this case is a direct short in the form of a copper wire.

With the switch open (Figure 24-5A) the applied voltage (E_a) will cause a voltage to appear across the primary winding. By transformer action a voltage will be induced in the secondary winding and will be indicated on the voltmeter as E_s. Assuming the transformer has a 1:1 turns ratio, E_s will be equal to the primary voltage. With no load there will be no reflected impedance, and the primary current will be determined by the applied voltage and the impedance of the primary winding; and the primary current will be relatively small.

An entirely different set of circumstances exist when the switch is closed (Figure 24-5B). The impedance of the secondary circuit becomes very small. This low secondary impedance (Z') is reflected back into the primary circuit in such a manner as to considerably lower the impedance of the primary circuit. Lowering

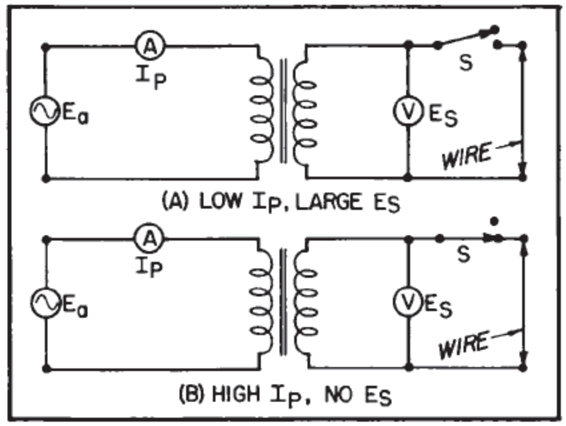


Figure 24-5 - Reflected impedance.

impedance of the primary circuit causes the primary current (I_p) to increase. Although voltage is still induced in the secondary winding, no secondary voltage (E_s) is measured by the voltmeter. This last statement brings up the second concept to be reestablished.

The secondary circuit can be thought of as a series circuit consisting of the induced voltage (acting as a source), the internal impedance of the secondary winding; and the zero resistance of the copper wire. Figure 24-6 shows a series equivalent of the secondary circuit. Notice that the entire induced voltage is dropped across the internal impedance of the winding, and no voltage is dropped across the copper wire.

NOTE: Voltage is still induced in the winding even though no voltage appears across the winding.

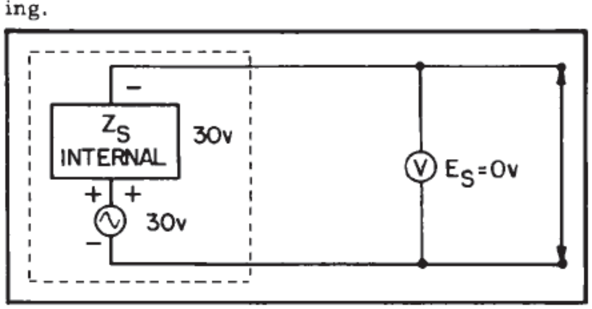


Figure 24-6 - Series equivalent of secondary circuit.

With these facts concerning a transformer firmly established, the theory of operation of the Armstrong oscillator shown in Figure 24-7 will now be discussed.

For ease of explanation it will be assumed that switch S_1 is opened, but the tube filament is hot and the tube is ready to conduct when S_1 is closed. A step by step analysis of a complete cycle of operation will now be discussed.

24-5. Initial Charge of Tank Capacitor

The conditions existing when S_1 is open are as follows:

Tube V_1 is ready to conduct. No current is flowing in the circuit, no fields around the coils, and no charge on the tank capacitor. While the following analysis will be drawn out step by step, it must be remembered that many of the actions are almost instantaneous and are occurring simultaneously.

1. Switch S_1 is closed. This applies plate voltage to V_1 and starts plate current flowing. Plate current will start flowing because there is zero potential on the grid at the first instant. The path for plate current is from the cathode through the tube, the plate load (L_1), the positive terminal of the power supply to ground; and back to the cathode of the tube. The plate current path is shown by the heavy arrows in Figure 24-7.

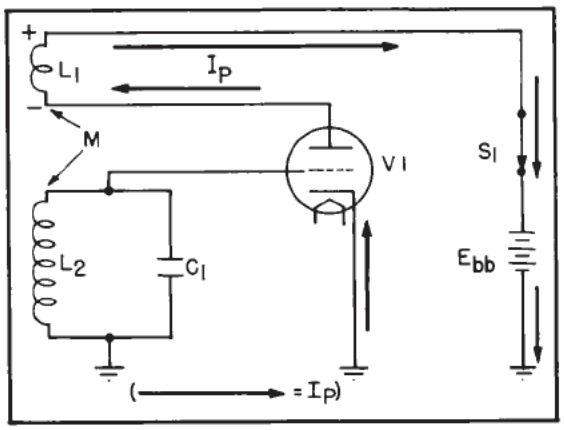


Figure 24-7 - Armstrong oscillator plate current path.

2. The dc resistance of the plate current path is very low, consisting only of the tube plate resistance and the small dc resistance of the plate coil L_1. At the instant S_1 is closed plate current will attempt to increase to its maximum Ohm's law value. At this instant plate current is experiencing a maximum rate of change. It would be expected that the maximum rate of change of plate current through the plate load would cause a large voltage drop across L_1. In actuality, however, the impedance of L_1 is very low at this time due to the reflected impedance of the tank circuit (to be explained). Therefore, the voltage dropped across L_1 at the first instant is relatively small because of the low load impedance. The polarity of this voltage drop is shown in Figure 24-8.

3. Due to mutual coupling between L_1 and L_2 the expanding field around L_1 will induce a EMF in L_2. The voltage induced in L_2 wil be maximum at the first instant, and due t the manner of connection of the coils in th circuit, will have the polarity shown i Figure 24-8.

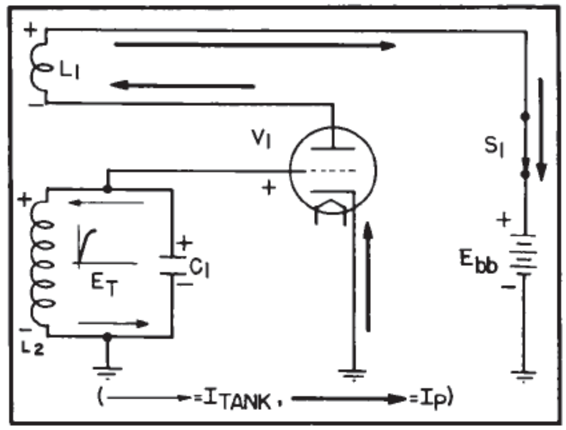


Figure 24-8 - Tank coil voltage.

However, while the voltage induced in L; will be relatively large, THERE WILL BI NO VOLTAGE ACROSS THE TANK AT THI: INSTANT. The truth of the above statemen can be seen if it is remembered that at th first instant of operation the tank capacito: has no charge, and therefore represents a short circuit to the circulating tank current Thus, the entire induced voltage is droppe across the internal impedance of tank coil L_2. Since the tank capacitor represents a short at the first instant the impedance o the tank will be extremely low. As the tanl capacitor assumes a charge the impedanc of the tank circuit will increase. It can nov be seen that the impedance reflected int plate coil L_1 by the tank circuit will be ver low at the first instant and then increase a: the tank circuit impedance increases.

4. The instant an induced voltage appears i tank coil L_2, capacitor C_1 will begin t charge using the induced voltage as a source. Thus, circulating current (I_{tank} begins.

   The voltage that appears across the capacitor is the tank voltage, and as such is th voltage applied between the grid and cathod of the tube V_1. As C_1 assumes a charg (Figure 24-8) the grid potential become: positive increasing plate current flow. Th voltage waveform (E_t) across the tank during the charge of C_1 is shown in Figure 24-8

5. The increased plate current flow result: in an increase in the density of the mag-

netic field of L_1. The increased density of the plate coil field causes an increase in the induced voltage of L_2. Therefore, C_1 will charge to a higher value again increasing the grid potential and the plate current.

6. This action is called POSITIVE or REGENERATIVE FEEDBACK and if the feedback is sufficient to overcome the losses of the circuit the plate current will continue to increase. Two factors prevent the increase of plate current to an indefinite value. The first is the non-linearity of the electron tube. The second is the increasing impedance of the plate load. As C_1 charges the impedance of the tank circuit increases. Since this impedance is reflected into the plate circuit, the plate circuit impedance must also increase.

7. The increasing plate load impedance and the increasing plate current causes a rapidly increasing voltage drop across the plate coil L_1. Plate voltage equals E_{bb} MINUS the drop across the plate load. Thus, a point is reached where the rate of change of plate current begins to decrease. This occurs because the plate voltage has been lowered to the point where the potential is no longer sufficient to attract all the ADDITIONAL electrons even though the grid voltage is increased. This is the non-linear operating region of the electron tube.

8. Since the rate of change of plate current decreases, the rate at which the magnetic field of L_1 cuts the turns of L_2 will also decrease. This results in a decrease in the amount of voltage induced in the tank coil.

Q2. What happens to the induced secondary voltage of a transformer when a load having zero resistance is placed across the secondary winding?

Q3. What is the charging current for C_1?

24-6. Discharge of Tank Capacitor

1. At the point where the induced voltage of L_2 falls below the magnitude of the charge on C_1 the tank capacitor will begin to discharge. As the tank voltage, and hence the grid potential, begins to decrease, the plate current will also begin to decrease.

2. Figure 24-9 shows the reversal of tank current due to C_1 discharge. The discharge of C_1 is slow at first because L_2 develops a CEMF almost equal to the charge of C_1. The collapse of the magnetic field of L_1 aids the discharge of C_1 by inducing a voltage in L_2 which opposes the CEMF in L_2 by the discharging tank capacitor. The polarity of L_2 remains the same because the tank current is the stronger of the two forces.

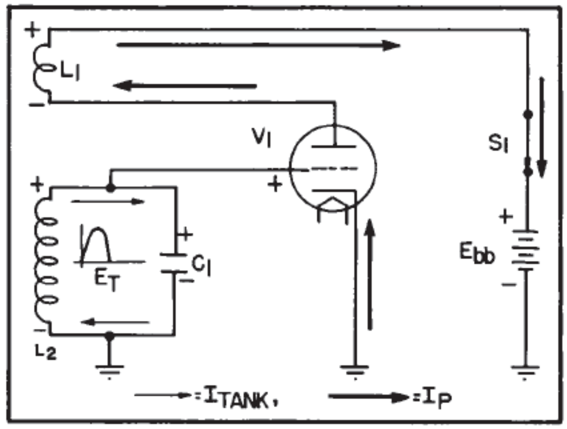


Figure 24-9 - Tank capacitor discharge.

3. As the tank capacitor discharges through the tank coil an increasing magnetic field is built up around L_2. As the charge on C_1 continues to decrease towards zero the grid potential also decreases toward zero. This, in turn, continues to decrease plate current. The regenerative feedback at this time is the collapsing of the plate coil's magnetic field aiding the discharge of C_1. The rate of change of the tank capacitor's discharge current INCREASES as the charge of C_1 approaches zero.

4. At the instant the charge of C_1 reaches zero the magnetic field around L_2 is maximum, the grid potential is zero, and plate current has approximately the same value as at the first instant of starting. The tank circuit has now completed one-half cycle of operation and has generated the voltage waveform (E_t) shown in Figure 24-9.

Q4. Would the value of the plate load impedance be high or low at the instant the charge of C_1 reached zero? Explain.

24-7. Polarity Reversal of Tank Voltage

1. A half cycle of operation has now been completed. The magnetic field of L_2 will begin to collapse maintaining tank current flow in the same direction. The collapse of L_2's field will begin charging C_1 with a polarity opposite to that of the first half cycle. As C_1 assumes a negative charge the grid potential will become increasingly negative.

2. The increasing negative potential on the grid will continue to decrease plate current. As plate current decreases the field of L_1 will continue to collapse, now aiding the charge of C_1 with the opposite polarity.

3. When the potential on the grid becomes sufficiently negative the tube will cut off and plate current will cease to flow. However, the field of L_2 will continue to collapse and tank current will continue to flow until C_1 is completely charged. Figure 24-10 shows the circuit conditions immediately after the tube has cut-off. Notice the tank current still flowing, the negative potential on the grid (due to the charge of C_1), and no plate current flowing (consequently no drop across L_1).

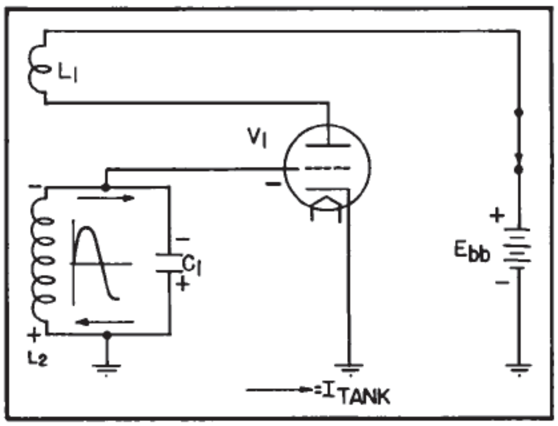


Figure 24-10 - Conditions at cut-off

4. At the instant C_1 is fully charged in the negative direction the field of L_2 will be zero, no tank current will flow, and the circuit will have completed three-quarters of a cycle of operation. The tank voltage waveform (E_t) for the first three-quarters of a cycle is shown in Figure 24-10.

Q5. What is the value of plate voltage at the instant C_1 is fully charged in the negative direction (Figure 24-10)?

24-8. C_1 Discharge on Negative Alternation

1. C_1 will now begin to discharge through L_2. As C_1 discharges the potential on the grid will become less negative. Tank current will be in the same direction as for the first quarter cycle and will build a magnetic field around L_2. During this period of time, energy is transferred from the electrostatic field of the capacitor to the electromagnetic field of the coil.

2. As the grid becomes less negative a point will be reached where the tube will com out of cut-off and plate current will begi to flow.

3. Plate current flow will again cause an ex panding field about L_1 which will induce voltage in L_2 of such a polarity as to ai the discharge of C_1.

4. When the charge on C_1 decreases to zer the circuit has completed a full cycle c operation and one complete sine wave c voltage has been generated across the tank An instant later the field about L_2 will be gin to collapse forcing current around th tank circuit in a counter-clockwise direct ion. The voltage induced in L_2 as a resu of the collapsing field thus begins to re charge C_1. Since the action at this instar is essentially the same as that which oc curred during the first quarter cycle (initia charge of C_1) of operation, it can be see that the entire action will be repeated an a second sine wave of voltage will be ger erated nearly identical to the first. If th plate and heater supply voltages are main tained the circuit will continue to generat an uninterrupted train of sine waves.

Notice that the sine wave is generated in th tank circuit and not by the tube. The tub functions as a switch which connects th power supply to the tank circuit at the pro per time to support oscillations. Since th sine wave is generated in the tank circu the frequency of the generated voltage i determined by the values of capacitance an inductance in the tank. The larger the valu of capacitance or inductance the lower wi be the frequency of the generated sine wave

24-9. Summary of Simple Armstrong Oscillato

1. When plate voltage is first applied to th oscillator, tube plate current will flow.

2. Plate current flow through the plate loa coil will cause an expanding magnetic fiel around this coil.

3. The expanding field of the plate load induce a voltage in the tank coil. The induced volt age causes the tank capacitor to charge.

4. The charge of the tank capacitor cause plate current to increase by increasing gri potential. This regenerative action con tinues until the non-linear characteristic of the tube cause a decrease in the rate c change of plate current.

A2. The entire induced voltage is dropped across the internal impedance of the winding.

A3. The circulating current, I_{tank}.

A4. Low. The value of reflected impedance would be very low, thus, lowering the plate load impedance.

A5. Equal to E_{bb} because no plate current is flowing.

---

5. When the induced voltage of L_2 falls below the charge of the capacitor, the tank capacitor begins to discharge. The discharge of C_1 causes the grid potential to decrease, thereby decreasing plate current.

6. When C_1 is completely discharged, the field of L_2 collapses, and charges C_1 with the opposite polarity. Part way through this portion of the cycle of operation the grid potential will become sufficiently negative to cut the tube off.

7. When the field of L_2 is completely collapsed, C_1 will begin to discharge. As the tank voltage becomes less negative the grid potential will approach the point where the tube will come out of cut-off. As the tube begins to conduct, regenerative feedback occurs and replaces the lost energy. Oscillations can now continue until dc power is removed from the circuit.

This circuit, however, has several disadvantages. The tube current varies between cut-off and its maximum limit. This produces a high average plate current and poor efficiency. Feedback is provided for all but a small portion of the 360° of the input and is of a high amplitude. This is unnecessary as the feedback amplitude and duration should only be great enough to overcome the tank I^2R losses. Since the grid is positive for a complete half cycle appreciable grid current will flow. This current flow is from the grid, through the tank, and back to the cathode thus, causing amplitude distortion of the output signal. Some form of bias is necessary to provide greater efficiency and definite control of feedback time. Methods of providing this bias are discussed in the following topics.

Q6. Why is the efficiency of the simple Armstrong oscillator so poor?

24-10. Classes of Operation

The class of operation used in an oscillator is determined mainly by the output requirements of the circuit. Class A operation is used when purity of the output waveform is desired and efficiency is unimportant. An audio oscillator used in test equipment is one example of this type of requirement.

Class C operation is generally used with RF oscillators. In this application the distortion of the output waveform is tolerable and greater efficiency can be realized. Since high Q tank circuits are used in RF applications, the conduction interval of the tube can be reduced to 90° or 120° of the 360° input cycle without adverse effects on the output waveform. This is possible because of the lower I^2R losses of the high Q tank circuits.

Q7. What change in plate dissipation occurs as the operation of an oscillator is changed from class A to class C?

24-11. Fixed Bias

One method of producing class C operation of an oscillator is by the use of fixed bias. An Armstrong oscillator using fixed bias is illustrated in Figure 24-11. The grid supply E_{cc} provides a negative bias that is equal to twice the cut-off potential of the tube. For example, if the tube cuts off with a grid potential of -30 volts then E_{cc} will supply a grid bias of -60 volts. A cycle of operation for the fixed bias Armstrong oscillator will now be analyzed. Assume the following circuit conditions to exist: tube filaments are at correct operating temperature, the grid is being held at the potential of E_{cc} (-60 V), plate voltage has not been applied.

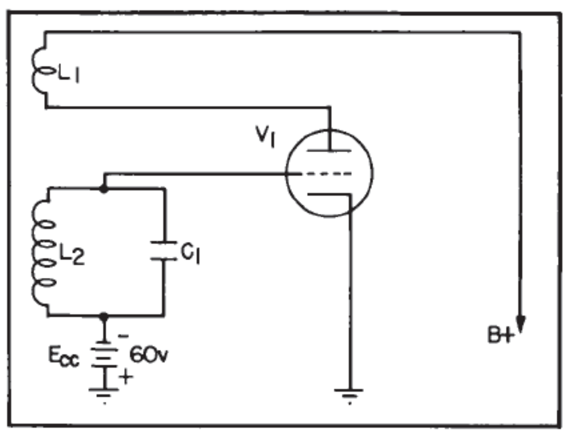


Figure 24-11 - Armstrong oscillator with fixed bias.

1. Upon application of plate voltage plate current can not flow because the tube is held at cut-off by grid bias. Unless the grid is

reduced in some manner the oscillator will not start.

2. Assume an external source is used to charge C_1 so that the grid end of the tank becomes positive. As the tank voltage becomes positive it will oppose E_{cc}. The tube cut-off voltage is -30 V. Thus, when the tank voltage becomes 30 V positive the tube will come out of cut-off.

   The circuit conditions at the instant the tube comes out of cut-off are shown in Figure 24-12. Notice that the tank voltage is 30 volts positive and e_c equals -30 V.

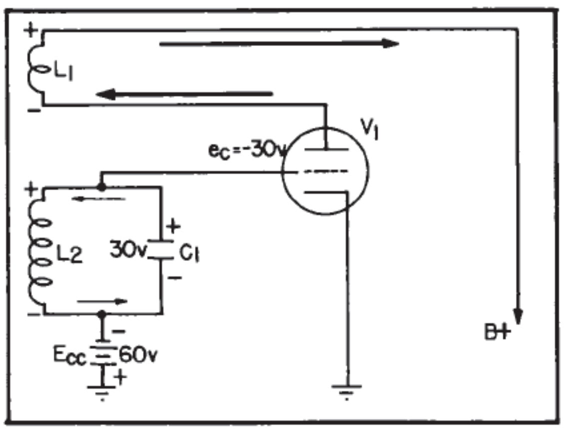


Figure 24-12 - Circuit conditions when tube begins to conduct.

3. As the tube comes out of cut-off, plate current will begin to flow (Figure 24-12). The action is now similar to that discussed previously for the Armstrong oscillator. The increase of plate current causes L_1 to induce a voltage in L_2, increasing the charge on C_1 (regenerative feedback).

4. As previously explained, the tube characteristics will cause the induced voltage to fall below C_1 charge. C_1 then begins to discharge and the tank voltage begins to decrease. This causes plate current to decrease.

5. When the positive tank voltage decreases enough to make the grid voltage more negative than cut-off potential, plate current will stop flowing. C_1 continues to discharge until the tank voltage is zero. At this instant:

$$e_c = E_{cc} = -60 \text{ V}$$

6. Flywheel effect causes the tank voltage to go through a negative alternation as C recharges (with the opposite polarity) ar discharges. During this period tank volt age is aiding E_{cc} and the tube remair cut-off.

7. When flywheel effect causes the positiv alternation of tank voltage to begin, th entire cycle repeats from step 2.

It can be seen that feedback occurs only dur ing a portion of the tanks positive alternatic (feedback occurs only when the tube is conduc ing). Because plate current takes the form of high amplitude, short duration pulse, the averaş plate current is low and circuit efficiency high.

The disadvantages of this method of supplyir bias for class C operation are: first, the circu is not self starting. An external source must t used to cause the initial feedback cycle. Se ond, if the tank voltage becomes insufficient amplitude to overcome the fixed bias, feec back will cease and the oscillations will dampe out.

GRID LEAK BIAS

Grid leak bias is used universally in trioc oscillators rather than fixed bias. The use grid leak bias makes the oscillator self star ing and provides more stable operation thi fixed bias. This type of bias is very suitab for an application where it is desired to ha zero bias under no signal conditions and negative bias when an input signal is applie This method of bias commonly appears in oı of two forms: SERIES or SHUNT grid leak bia Before entering into a discussion of the opera ion of oscillator circuits employing grid le bias, a brief review of RC time constants wi prove advantageous. A reader desiring a mo: comprehensive review is directed to Chapter 1

24-12. Charge Time Constant

If a potential is applied across a capacito the capacitor will charge almost instantaneous to the value of the applied potential. Howeve if a circuit contains resistance in series with tl capacitor the charge time will be altered. Tl time involved for a series resistor and capacit combination to reach 63.2% of its maximu charge is called the TIME CONSTANT (T) the circuit. The mathematical relationship the resistance, capacitance, and time consta is expressed in the equation:

$$T = RC \qquad (10\text{-}1$$

A6. Grid current flows for a full half cycle, raising the amount of power dissipated in the input circuit.

A7. Plate dissipation decreased due to the decreased conduction angle and lower average value of plate current.

---

where: T = one time constant in seconds
R = circuit resistance in ohms
C = circuit capacitance in farads

For most practical applications a capacitor can be assumed to have reached full charge after five time constants (5T) have elapsed.

The significance of the RC charge time constant, as regards grid leak bias and oscillators, may be examined by use of the partial circuit in Figure 24-13. The circuit consists of a coil (L) (which is actually the secondary winding of a transformer), and a 100 pf capacitor (C) which is connected in series with the parallel combination of a 1 meg resistor (R) and the input resistance of the tube (grid to cathode resistance).

Assume that a potential is applied to (L) having the polarity as shown in Figure 24-13. Current will be caused to flow in the direction indicated. Since the capacitor appears as a short at the first instant the full potential of the source (L) will appear across the parallel combination of resistor (R) and the grid to cathode resistance of the tube. This voltage drop will cause the grid to assume a positive potential and draw grid

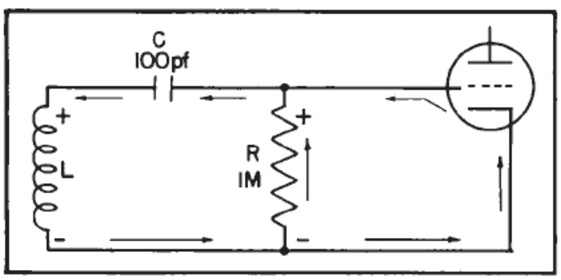


Figure 24-13 - RC charge time.

current. The magnitude of the grid to cathode resistance (while grid current is being drawn) may be on the order of 1K ohm. Since R is 1 meg ohm the total resistance of the parallel combination is, for all practical purposes, 1K ohm. The time constant of this circuit during charge would then be:

$$T = RC$$

$$T = (1 \times 10^3)\times(100\times10^{-12})$$

$$T = 1 \times 10^{-7} \text{ seconds}$$

The capacitor will assume a full charge in 5 time constants. Therefore, the full charge time will be:

$$\text{full charge time} = 5T$$

$$\text{full charge time} = 5\times10^{-7} \text{ seconds}$$

Thus, the capacitor will assume a full charge in approximately one half a microsecond. From the preceeding it can be seen that if a sine wave (up to 500 kc) were applied to L, the rise of the charge on the capacitor would very nearly follow the rise of voltage across L. Figure 24-14 shows the circuit with a sine wave input signal.

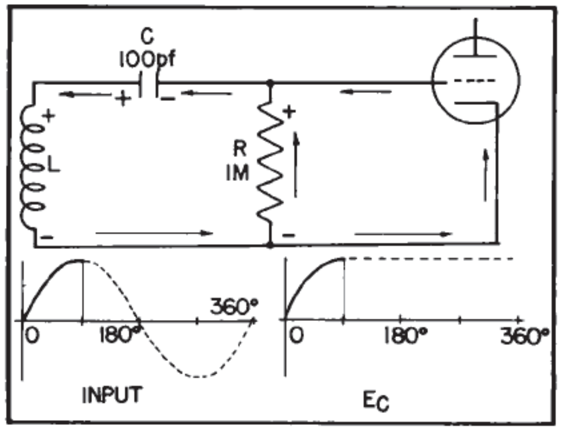


Figure 24-14 - Sine wave input (0° to 90°) capacitor charge.

From zero degrees to 90 degrees of the input signal the capacitor is charging. The current path and polarities are as shown. As stated, the charge time is so small as to allow the capacitor charge to follow the voltage rise.

24-13. Discharge Time Constant

Figure 24-15 shows the conditions existing in the circuit between 90° and 360° of the input signal. It is assumed that the capacitor charge reached maximum at the same time the input signal reached its positive peak (90°). As the input signal passes 90° and begins to become less positive the potential on L becomes less than the charge on the capacitor. The capacitor will begin to discharge. However, since the grid does not emit electrons, the low resistance path between grid and cathode does not exist for the discharge current. Instead, the path for capacitor discharge current is through the 1 meg, resistor, transformer winding L, and back to the capacitor. The time constant of the circuit during discharge is:

$$T = RC$$

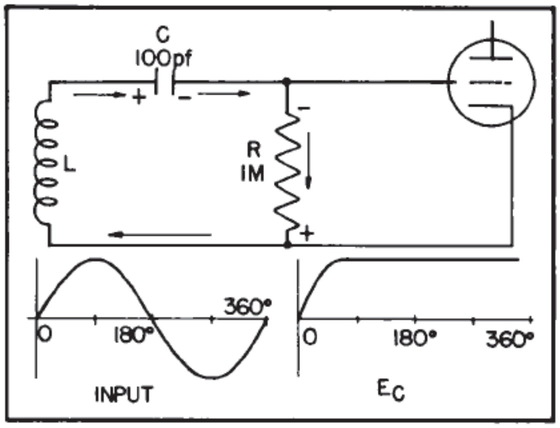


Figure 24-15 - Sine wave input 90° to 360° capacitor discharge

$$T = (1\times10^6)\times(100\times10^{-12})$$

$$T = 1 \times 10^{-4} \text{ seconds}$$

The time required to discharge C is now:

full discharge time = 5T

full discharge time = 5×10^{-4} seconds

Thus, the charge time of the circuit is a half of a microsecond while the discharge time is 500 microseconds. If the frequency of the input signal were 500 Kc, one cycle would take approximately 2 microseconds. The capacitor will be unable to discharge a significant amount before the next positive peak. The graph of voltage across the capacitor (Figure 24-15) shows that the capacitor will rapidly charge to the peak value of the input signal and maintain this value. NOTE: At this point in the explanation the capacitor is assumed to reach full charge within the space of one cycle. In a practical circuit the capacitor will take many cycles to acquire a full charge.

The attempted discharge of C through the grid resistor R applies a negative potential to the grid (notice polarity of the resistive voltage drop in Figure 24-15).

Q8. If a circuit, such as the one in Figure 24-15, contained an 80 pf capacitor and a 1.5 meg. resistor, what would the time constant (T) be for charge? For discharge? Assume the grid to cathode resistance of the tube as 1K ohm and infinity respectively.

24-14. Development of Grid Leak Bias

Bias is defined as the AVERAGE value of dc voltage between the grid and cathode of a tube. If the average potential between grid and cathode of a tube is zero, the bias will be zero. Application of a sine wave will cause the instantaneous value of grid voltage to vary but the average grid voltage, or bias, will remain zero.

This is shown by use of the partial schematic in Figure 24-16. Part A of the figure shows the input circuit of a tube with a coil connected between grid and cathode. The coil is the secondary of a transformer; thereby, permitting an external signal to be coupled into the circuit. A 20 volt peak signal is applied across the coil. Part B of Figure 24-16 illustrates a plot of the voltage between grid and cathode. Notice that the average voltage is zero even though the instantaneous grid voltage varies between plus 20 and minus 20 volts.

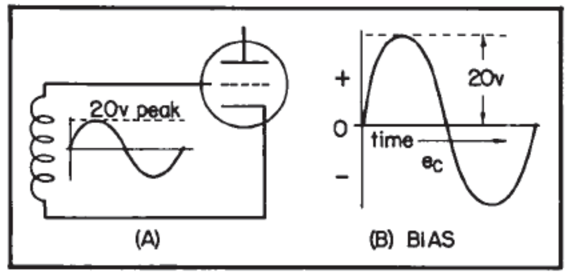


Figure 24-16 - Zero bias.

The addition of a resistor and capacitor in series with the coil will alter the bias. Figure 24-17 illustrates the circuit with a series resistive-capacitive network added.

It was previously shown that the capacitor will charge rapidly to the magnitude of the applied voltage and (for all practical purposes) maintain this charge as long as the magnitude of the applied voltage is not changed. Thus, application of a 20 volts peak sine wave to the coil in Figure 24-17 will cause the capacitor to charge to 20 volts with the polarity as shown. The capacitor will attempt to discharge through the large value grid resistor, and thereby, produce a negative potential, -20V, on the grid of the tube.

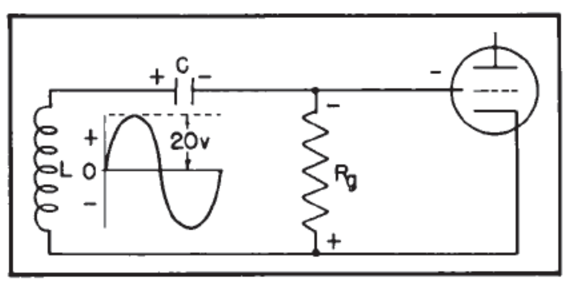


Figure 24-17 - Negative bias.

In order to better understand the operation of the bias circuit when a signal is applied, Figure

A8. Charge: T = 0.08 microseconds.
Discharge: T = 120 microseconds.

24-17 is redrawn as an equivalent circuit in Figure 24-18A. Since the voltage between grid and cathode is the same as that across R_g, the output of the equivalent circuit is taken across R_g. The 20 volts across the capacitor is assumed to stay relatively constant while the voltage across the coil varies from plus 20V to minus 20V. The output voltage, e_c, is graphed in part B of Figure 24-18.

At the instant of time zero (t_0) the input signal is zero. The output voltage, e_c, is the algebraic sum of the capacitor voltage and the coil voltage. Since the coil voltage is zero at t_0, the only voltage present across resistor R_g is the -20V from the capacitor. Thus, at t_0 the instantaneous grid voltage (e_c in Figure 24-18B) is -20 volts. As the input signal increases in the positive direction a voltage is developed across the coil, with a polarity of positive at the top (capacitor end) of the coil and negative at the bottom. The polarity of the coil voltage (during the positive half cycle of the input signal) is such as to subtract from the capacitor voltage.

At t_1 the capacitor voltage is still -20V, but the input signal has now developed 20V of the opposite polarity across the coil. Thus, at t_1 the voltage (e_c) across R_g is:

$$(-20) + (20) = 0 \text{ volts}$$

At t_2 the input signal is again passing through zero. There is no voltage present across L, and e_c is again equal to the capacitor voltage (-20V).

As the input signal goes through its negative half cycle the polarity of the induced voltage in L is such as to aid the capacitor voltage. Therefore, when the input signal attains its maximum negative peak of -20 volts at t_3, e_c will be:

$$(-20) + (-20) = -40 \text{ volts}$$

At t_4 e_c will again equal -20 volts.

It can be seen from the above that the addition of the resistor and capacitor combination has caused the average voltage between grid and cathode (bias voltage) to become -20 volts.

It has been assumed, during this explanation, that the capacitor was capable of maintaining a constant voltage. In actuality the capacitor will lose a small percentage of its charge between the positive peaks of the input signal. This lost charge is replaced by driving the grid SLIGHTLY positive for a brief interval during each positive peak of the input signal. For instance, assume that between t_1 and t_4 (Figure 24-18B) the capacitor voltage decreases to -19 volts. The positive peak of the next cycle of the input signal will cause L to develop 20V. The instan-

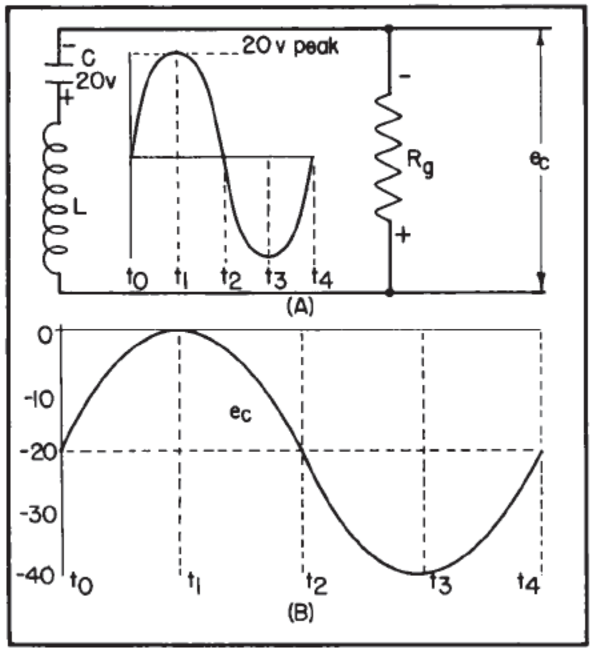


Figure 24-18 - Equivalent circuit.

taneous grid voltage e_c will then be:

$$(-19) + (20) = 1 \text{ volt}$$

During the interval the grid is driven 1 volt positive; grid current will be drawn, and the capacitor will again be charged to 20 volts (as explained in section 24-12).

Therefore, the bias voltage will continue to be approximately -20V as long as the magnitude of the input signal does not vary. If the input signal were to increase in magnitude, the capacitor (in a few cycles) would charge to the new value and then maintain it as the bias voltage. On the other hand, if the input signal were to decrease in magnitude the capacitor charge would decrease (in a few cycles) to the new value and maintain it as the bias.

Under proper operation grid leak bias will maintain bias relatively near the peak value of the input signal.

24-15. Shunt Grid Leak Bias

Figure 24-19 shows an Armstrong oscillator using SHUNT grid leak bias. The action of the oscillator and the grid leak components has been discussed in detail in the previous sections.

When the tank voltage becomes more positive than the charge on C_{g1}, grid current will be drawn and C_{g1} will charge. When tank voltage

is less positive than the charge on C_{gl} the grid leak capacitor will discharge through the resistor R_{gl} causing a negative potential to be applied to the grid.

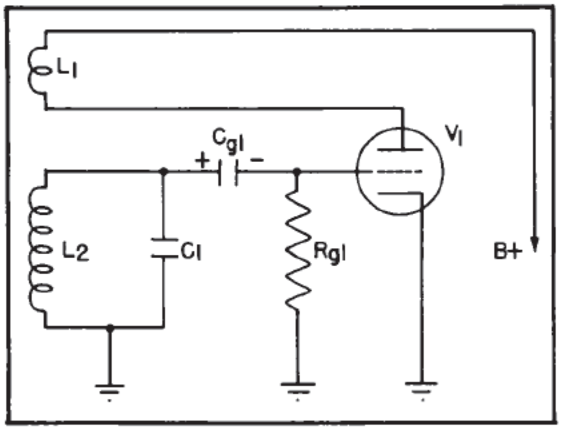


Figure 24-19 - Armstrong oscillator with shunt grid leak bias.

Shunt grid leak bias derives its name from the fact that the grid leak resistor is in parallel, or shunt, with the tank circuit.

24-16. Series Grid Leak Bias

Figure 24-20 illustrates an Armstrong oscillator with SERIES grid leak bias. The operation is similar to the shunt type, differing only in discharge path of C_{gl}. The charge and discharge paths are shown in Figure 24-20. Series grid leak bias is so named because the grid leak resistor is in series with the tank circuit.

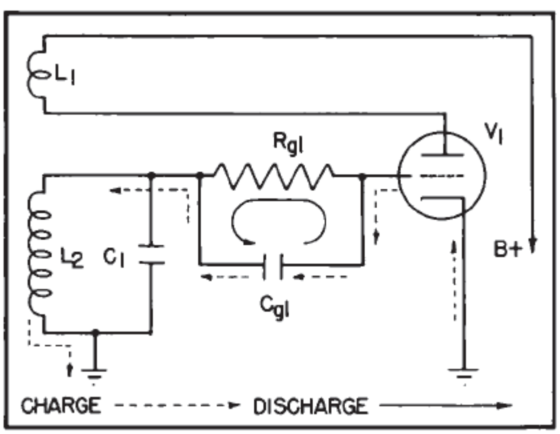


Figure 24-20 - Armstrong oscillator with series grid leak bias.

One of the advantages of grid leak bias for oscillators is that the circuit is self starting. This is possible because when the circuit is first energized C_{gl} has no charge. Therefore, bias is zero and plate current can flow. This will initiate feedback and start oscillations. Anothe advantage is the circuit is self adjusting. Thi is due to the bias level being dependent on th input signal amplitude. One disadvantage of gri leak bias comes from the zero bias with no ap plied signal. Should a malfunction occur in th tank circuit or bias network, the bias will be come zero, causing a steady flow of large plat current until the circuit is deenergized.

Q9. Why does grid leak bias improve the ef ficiency of the Armstrong oscillator?

Q10. In an oscillator using grid leak bias doe plate current only flow when the grid is positive Explain.

Q11. Will an oscillator work if it is not oper ated class C? Explain.

OUTPUT COUPLING METHODS

A circuit which will produce stable, self sustained oscillations is of no value unless ther is some way to couple a useful output from i There are two methods of coupling in genera use. They are INDUCTIVE COUPLING an CAPACITIVE COUPLING.

24-17. Inductive Coupling.

INDUCTIVE coupling is illustrated in Figur 24-21. In this case the output is taken acros L_o which is inductively coupled to the tank co

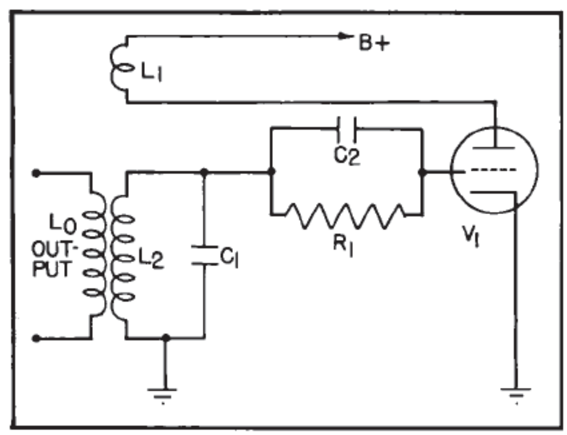


Figure 24-21 - Inductive coupling (output).

L_2. In some cases L_o is shunted by a capacito to form a tank circuit which is resonant to th f_o of the oscillator. This is known as a tune secondary.

In either case, the output amplitude is de termined by the degree of coupling between L and L_2. If L_o and L_2 are physically close, th

A9. It normally produces class C operation thus improving the efficiency.

A10. No. Plate current flows as long as the bias is above cut-off.

A11. Yes. The class of operation merely determines the efficiency of the circuit.

---

output amplitude is high, and a considerable amount of power is transferred from the oscillator tank. This is known as tight coupling which has the disadvantage of reflecting a high I^2R loss back to the tank circuit, effectively lowering tank Q. If the distance between L_o and L_2 is increased, less power is transferred from the oscillator tank, and the output amplitude decreases. This is known as loose coupling. Loose coupling is usually used because the reflected losses are low and tank Q is not appreciably affected. This results in a greater degree of amplitude and frequency stability. The low amplitude of the output signal presents no problem as amplification is easily accomplished.

24-18. Capacitive Coupling

CAPACITIVE coupling is illustrated in Figure 24-22. The coupling capacitor (C_C) is of a low value in order to present a high X_c at the resonance frequency. With this method the tank is shunted by a high value of reactance, which remains constant. The effect of this shunt reactance can be compensated for by slight variations of the tank L or C values.

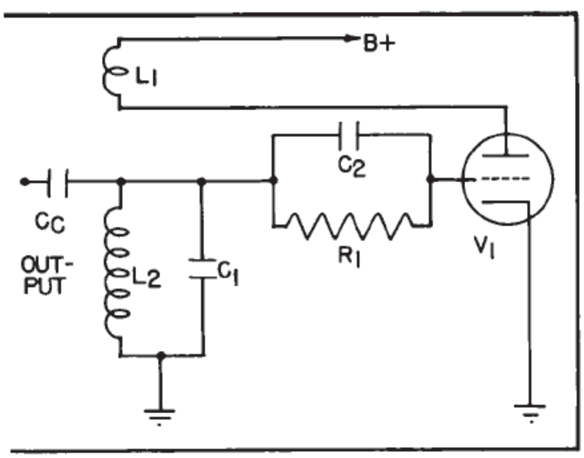


Figure 24-22 - Capacitive coupling (output).

Q12. If the value of C_c (Figure 24-22) were increased what would be the effect on the tank circuit Q?

STABILITY

24-19. Amplitude Stability

AMPLITUDE STABILITY can be affected by changes in bias, tube gm, supply voltage, and reflected impedance.

An increase in bias would cause a decrease in feedback time and amplitude. The decreased feedback would lower the amplitude of the oscillations. A decrease in bias would produce the opposite effects.

A decrease in the gm of the amplifier would reduce the gain of the circuit and produce the same effect as increased bias. A change in E_{bb} will shift the operating point of the circuit, consequently changing the amplitude of oscillation.

If the impedance of the load decreases it will reflect an increased I^2R loss back to the tank. If the feedback amplitude remains constant this loss will not be compensated for and the amplitude of oscillation will decrease.

Amplitude stability can be improved by using grid leak bias. The improvement is due to the bias level being dependent on the amplitude of the tank signal.

24-20. Frequency Stability

FREQUENCY STABILITY can be affected by changes in temperature, tank Q, reflected I^2R losses, supply voltage, and vibration.

Changes in temperature will cause the tube interelectrode capacitances to vary. It will also cause slight variations in the values of tank L and C. These changes will shift the output frequency. Since these changes occur slowly, the resultant shift in frequency is called DRIFT. Frequency drift can be minimized by adequate ventilation, circuit components which can easily dissipate heat, or in some cases, a temperature control system.

Mechanical vibration can cause the same variations in component values as temperature changes did. In this case, however, the resultant frequency shift is more rapid in nature. The oscillator chassis is often shock mounted to overcome this problem.

Variations in supply voltage will change the operating point of the circuit and cause circuit instability. This can be overcome by use of a regulated power supply.

The load impedance, reflected into the tank circuit, has a great effect upon the frequency stability of the oscillator. The reflected impedance may contain both resistive and reactive components. The resistive component will lower the Q of the tank circuit, and the reactive component will alter the resonant frequency. When the Q of the tank is very high and the reflected impedance is small, the effect on the resonant frequency and tank Q is negligible. The lower the effective Q of the tank the greater will be the ef-

fect of a change in load impedance on the resonant frequency. High Q tanks and loose coupling reduce the frequency instability caused by these factors.

Q13. Could the frequency stability of the oscillator ever exceed the frequency stability of the tank circuit?

OSCILLATOR CIRCUITS

A number of other oscillator circuits exist in addition to the Armstrong oscillator. In many cases these oscillators are named after the men who devised them. Thus, an oscillator may be classified as an Armstrong, Hartley, Colpitts, etc.

Another way of classifying an oscillator circuit is based on the manner in which dc power is applied to the tube. An oscillator in which dc power is supplied to the tube through the tank circuit, or a portion of the tank circuit, is called SERIES FED. Any oscillator in which dc plate current flows through the tank or a portion thereof is series fed.

An oscillator which receives its dc plate power through a path separate and parallel to the tank circuit is said to be parallel or SHUNT FED. The tank circuit of a shunt fed oscillator does not carry any dc plate current. All of the oscillators to be discussed in this Chapter can be either series or shunt fed.

24-21. Series Fed Hartley Oscillator

The SERIES fed Hartley oscillator circuit is illustrated in Figure 24-23.

The tank circuit is composed of C_1, L_1, and L_2. L_2 also provides a means of supplying feedback to the tank circuit. Grid leak bias is provided by C_2 and R_1. The RFC (radio frequency choke) and C_3 provide plate decoupling and prevent the RF variations from appearing across the power supply.

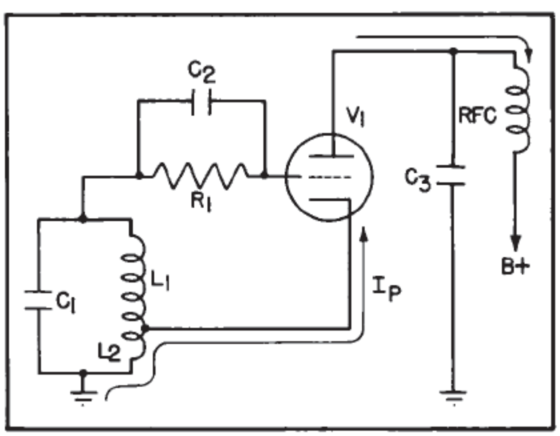


Figure 24-23 - Series fed Hartley oscillator.

The operation of the Hartley oscillator is similar to that of the Armstrong; the main difference being the method of feedback. In the case of the Hartley the feedback coil L_2 is part of the tank circuit.

The first cycle of operation starts with the initial increase in plate current which is caused by the application of plate voltage. Grid bias is equal to zero. Since L_2 is in series with V_1, plate current flow through L_2 will produce a voltage drop across it. The expanding field of L_2 will induce a voltage in L_1 in such a manner as to make the top of L_1 positive. C_1 will begin to charge causing the tank voltage to increase in a positive direction. This places a positive potential on the grid of V_1, causing plate current to increase and drawing grid current to charge C_2. The cycle of operation continues in the same manner as for the Armstrong oscillator. As the amplitude of the oscillations build up, the bias will increase until the circuit stabilizes at class C operation. Plate current will take the form of pulses with the same frequency as the tank oscillations. The RFC presents a high reactance to this frequency blocking the pulses from the power supply. C_3 presents a low reactance to this frequency, thus, bypassing the pulses around the power supply. With this method of decoupling, RF variations are prevented from appearing across the power supply.

The main disadvantage of the series fed Hartley circuit is the fact that the dc tube current flows through part of the tank coil. This not only necessitates this section of the tank coil being designed to carry the heavier current but will also adversely affect the stability of the oscillator.

Q14. Starting at the cathode, give the path for the grid leak capacitor C_2 charging current.

Q15. Does current flow through L_2 during the time the tube is cut off? Explain.

24-22. Shunt Fed Hartley Oscillator

A variation in the method of component connection can overcome the disadvantage of plate current flowing through part of the tank circuit. The variation in connection is shown in Figure 24-24. When connected in this manner the circuit is known as a SHUNT fed Hartley oscillator. This circuit operates in the same manner as the series fed Hartley. The only difference being in the manner in which feedback is accomplished. The feedback cycle occurs in the following manner:

1. Assume that the oscillator is in operation, that the tank voltage is at the zero point of its cycle, and that the positive alter-

A12. Tank circuit Q would be lowered.

A13. No.

A14. From the cathode to the grid, through C_2, through L_1, and back to the cathode.

A15. Yes. The circulating current of the tank.

nation is about to begin.

2. At this time the tube is cut-off by the action of the grid leak bias and plate voltage is equal to B+. C_3 is effectively connected in parallel with the tube. Therefore, the charge on C_3 will be equal to the plate voltage of V_1.

3. During the positive alternation of the tank voltage the grid-to-cathode potential becomes less negative. Plate current flows when this negative potential becomes less than that required for cut-off. The increase in plate current causes a decrease in plate voltage (due to the action of the RFC).

4. Since the value of plate voltage is now less than B+, C_3 will start discharging through L_2. The voltage drop developed across L_2 is regeneratively coupled to L_1, increasing the tank voltage, causing the grid voltage to be even less negative.

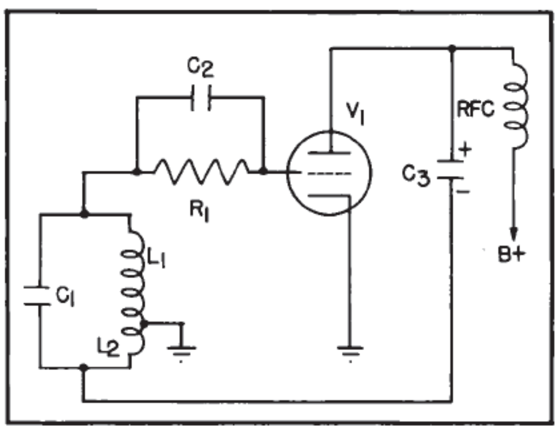


Figure 24-24 - Shunt fed Hartley oscillator.

5. Due to flywheel action, the tank voltage becomes less positive. Plate current decreases and plate voltage increases. C_3 now recharged to the increased value of plate voltage.

6. The charge current of C_3 flows through L_2 causing a voltage drop which aids the flywheel action and further reduces the tank voltage.

7. When the positive tank voltage decreases to a point where the instantaneous grid-to-cathode potential becomes negative enough to cut the tube off, plate current stops and the plate voltages rise to B+ potential. C_3 charges to plate potential and the feedback cycle stops. The tank continues to oscillate due to flywheel action.

In addition to providing decoupling, the RFC acts as a plate load for V_1 developing the variations of plate voltage which C_3 couples back to the tank. The size of C_3 will play a part in determining the amplitude of the feedback. If the size of C_3 is increased, X_{C3} will be lower and feedback will increase. If the size of C_3 is decreased, X_{C3} will be higher and the feedback will decrease.

It can be seen (Figure 24-24) that the path for plate current is from the cathode to the plate, through RFC, through the power supply, and back to the cathode of V_1. Thus, the plate current does not flow through any part of the tank.

Q16. Starting at the negative terminal of C_3, give the discharge path of C_3. (Figure 24-24)

24-23. Colpitts Oscillator

The COLPITTS oscillators, which are illustrated in Figure 24-25, are similar to the Hartley oscillators.

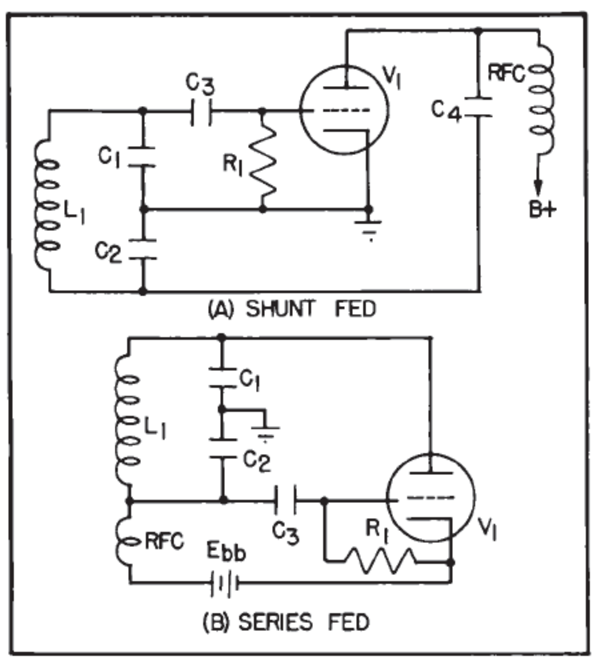


Figure 24-25 - Colpitts oscillators.

The main difference is in the type of feedback used. The Colpitts uses a split tank capacitance rather than a tapped inductor. In Figure 24-25A capacitors C_1 and C_2 act as a capacitive voltage divider across the tank circuit. Since the voltage across C_1 is the input signal to the tube the amount of feedback depends on the ratio between the capacitances of C_1 and C_2. If the capacitance of C_2 is decreased the reactance of C_2 will increase. This will cause more of the total tank voltage to drop across C_2 and less across C_1. Since the voltage across C_1 has decreased, the input signal to the grid is less; and greater feedback occurs.

The tank circuit is composed of L_1, C_1, and C_2. The feedback path, from plate to tank, is through C_4 and C_2. The RFC provides decoupling and develops the changes in plate voltage necessary for feedback. C_3 and R_1 form a shunt grid leak bias network. Shunt grid leak bias must be used to provide a dc path from grid to cathode. In the case of the Hartley circuit, the feedback was applied to a part of the tank inductor. In the Colpitts, the feedback will increase the charge on the tank capacitor. In the Hartley and Armstrong oscillators, the first cycle of oscillation started with the initial surge of tube current. In the Colpitts, the first cycle starts with the collapse of the magnetic field of tank inductor L_1. The operation of the series fed Colpitts (Figure 24-25B) is similar to the shunt fed. It is seldom encountered, so detailed analysis will not be presented in this chapter.

In order to examine a cycle of operation, assume that V_1's filament is warm and that plate voltage has not been applied. (Figure 24-26)

1. Plate voltage is applied to V_1. Since bias is zero plate current will start to flow. C_1, C_2, and C_4 will begin to charge. The path of charging current and polarities of capacitor charge are shown in Figure 24-26.

2. The charging current produces a magnetic field around L_1. When the capacitors complete their charge, charging current will stop.

3. The magnetic field of L_1 will start collapsing in an attempt to maintain current flow. The current caused by the collapsing field of L_1 is now the circulating current of the tank and its path is through C_1, C_2 and back to the top of L_1. The collapsing magnetic field and the discharge of C_2 increase the positive potential on the grid side of C_1. Tank voltage is now starting its positive alternation.

4. The positive tank voltage causes plate current to increase and plate voltage to decrease. The decrease of plate voltage causes C_4 to discharge into C_2 further increasing the difference in potential across the tank.

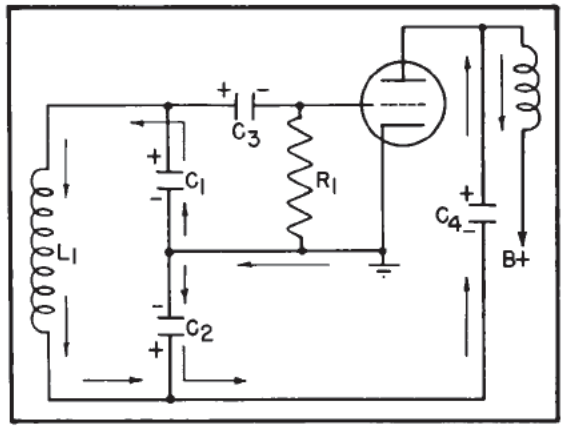


Figure 24-26 - Capacitor charging current for Colpitts oscillator.

5. This action will continue until the field around L_1 has completely collapsed. At this point tank voltage and plate current will stop increasing, and plate voltage will stop decreasing. The tank voltage is now at the positive peak of its first cycle.

6. Capacitors C_1 and C_2 start to discharge. Circulating current reverses direction. The positive tank voltage begins to decrease causing plate current to decrease and plate voltage to increase. C_4 starts to charge aiding tank action.

7. This action will continue until the tube is cut-off by grid leak action. At this time plate voltage is equal to B+, C_4 is no longer charging; and the feedback cycle has stopped. Tank oscillation is now maintained by flywheel effect.

Note the similarity between this circuit and the shunt fed Hartley in the method of feedback. In both circuits, feedback is produced when a change in plate voltage is coupled through a capacitor to the tank circuit. The major difference lies in the tank component to which the feedback is applied.

Q17. Why must a dc return path from grid to cathode be provided in the Colpitts oscillator?

24-24. Electron-Coupled Oscillator

The stability of the oscillators described in the previous sections is affected by the appli-

A16. Through L_2 to ground, from cathode to plate of V_1, and back to the positive terminal of C_3.

A17. Contact electrons sticking to the grid need a "leakage" path or they will eventually lower the gain of the tube and even possibly cut the tube off.

---

cation of a varying load. In order to lessen this effect the ELECTRON-COUPLED oscillator (ECO) is often used in applications where a varying load is encountered. The ECO actually consists of two separate circuits, the oscillator circuit and the output circuit. A single tube is used and the only coupling between the oscillator and output circuits is the electron stream of the tube. Hence, the name electron-coupled oscillator. One type of ECO is illustrated in Figure 24-27. The following breakdown will treat the oscillator section and the output section of the ECO separately.

The oscillator section is a shunt fed Hartley and is composed of the control grid and screen grid circuits of V_1. The screen grid serves as the plate of the oscillator section. R_1 and C_2 provide grid leak bias. R_2 is used to develop changes in the screen voltage and is sometimes replaced by RFC. C_3 provides a path for feedback from the screen grid to the tank circuit. Oscillations are produced in the same manner as in the shunt fed Hartley circuit previously described. Although the shunt fed Hartley is used in this explanation, several other types can be used in the oscillator section of the ECO.

The output section consists of the plate circuit of V_1. L_3 is the plate load of V_1. C_5 acts as a power supply bypass. The output is taken from the plate circuit by using either inductive or capacitive coupling.

The grid circuit of the ECO functions as a conventional triode oscillator. The control grid signal produces pulses of current in both the screen grid and plate circuits. The pulses of screen current are fed back to the grid tank to maintain oscillation. The plate current pulses cause plate voltage variations which are coupled to the output.

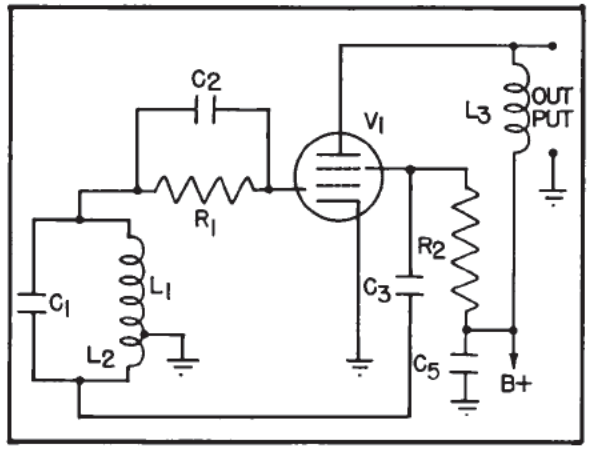


Figure 24-27 - Electron-coupled oscillator

Although the output circuit is dependent on the oscillator, the oscillator is INDEPENDENT of the output circuit. For instance, if the plate circuit were disconnected, the oscillator would still function normally. Therefore, variations in load have little effect upon the oscillator stability.

It should be noted that the plate current of the ECO occurs in narrow high amplitude pulses. When an RF choke is used as a plate load impedance, the plate voltage waveform consists of negative going spikes of voltage occurring at the same frequency as the sine wave developed in the tank circuit.

If a sine wave output is desired, an additional tank circuit is placed in series with the plate to act as a plate load. The plate current pulses will deliver energy to this tank causing it to oscillate and produce a sine wave. For proper operation, the tank must be tuned to the operating frequency of the oscillator, or to one of its harmonic frequencies.

EXERCISE 24

1. What does frequency stability of an oscillator mean?
2. What does amplitude stability of an oscillator mean?
3. What causes damped oscillations in an LC tank circuit?
4. What are the requirements to prevent damped oscillations?
5. What determines the rate of damped oscillations in an LC tank?
6. What is meant by the term "Flywheel effect"?
7. How can the resonant frequency of a tank circuit be made variable?
8. When the voltage across an LC tank circuit is decreasing, what is happening to the energy stored by the capacitor? By the inductor?
9. How many times does the capacitor in a tank circuit completely discharge during one cycle of operation?
10. What method of feedback coupling is used in the Armstrong oscillator?
11. Why is bias used in the Armstrong oscillator?
12. If cut-off in an Armstrong oscillator occurs at -10 volts and is operating class C, what should the bias be?
13. Why must the Armstrong oscillator be started by an external source when fixed bias is used for class C operation?
14. What methods of bias would be used for a self-starting Armstrong oscillator?
15. What would happen to the grid leak bias in the Armstrong oscillator if the feedback were increased?
16. What are the advantages of grid leak bias in oscillator circuits? Give detailed explanations of each.
17. What is usually sacrificed in an oscillato circuit in order to obtain a high-purit waveform?
18. Why would tight inductive coupling be use for the output of an oscillator?
19. Why would loose inductive coupling be use for the output of an oscillator?
20. When capacitive coupling is used for coupling from oscillator to output, what mus be the reactance of the capacitor? Why'
21. Why will a change in temperature affect a oscillator's frequency?
22. What should be the relative size of the loa impedance on an oscillator for maximun frequency stability? Why?
23. What type of feedback is used in the serie Hartley? In the shunt Hartley?
24. What is the main disadvantage of the serie fed Hartley oscillator?
25. What is the purpose of the RFC in the plat circuit of the shunt fed Hartley?
26. Are the series and shunt fed Hartley oscillators self-starting? Why?
27. Explain the method of feedback in the Colpitts oscillator.
28. What would be the effect on circuit operatio if series grid leak bias were used in th Colpitts oscillator (Figure 24-26)?
29. Which components primarily determine th resonant frequency of the Colpitts oscillator in Figure 24-26?
30. What would be the effects on circuit operation if C_4 opened in the Colpitts oscillato illustrated in Figure 24-26?
31. Which tube element in the ECO serves th same purpose as the plate in the Hartle oscillator as far as feedback is concerned?
32. What is one advantage of using an electron-coupled oscillator?

CHAPTER 25

RADIO FREQUENCY POWER AMPLIFIERS

A transmitter is a device for converting intelligence, such as voice or code, into electrical impulses for transmission through space from a radiating antenna. The simplest transmitter consists of an oscillator which generates a high frequency radio signal and an antenna system to propagate the energy. There are two main drawbacks to connecting the oscillator directly to the antenna. The first is that the power output would be limited because there are no stages of RF amplification between the oscillator and the antenna to build up the strength of the RF signal. Power output is important because it determines the distance over which the transmitted signal can be picked up by a receiver. The other consideration is frequency stability. The load impedance of the oscillator, in this case the antenna, reflected into the tank circuit has a great effect upon the frequency stability of the oscillator. The reflected impedance may contain both resistive and reactive components. The resistive component will lower the Q of the tank circuit and the reactive component will alter the resonant frequency. A drift in the frequency of the transmitted signal would mean that a portion of the message would be lost by the operator trying to receive it. To overcome the limitations of connecting an oscillator directly to the transmitting antenna, one or more stages of amplification are connected between the oscillator and the antenna. These stages are called RF POWER AMPLIFIERS. The stage which is connected to the antenna is usually called the FINAL POWER AMPLIFIER. The other stages of amplification are known as INTERMEDIATE POWER AMPLIFIERS (IPA). The first power amplifier, since it serves to isolate the oscillator from variations of load, is called a BUFFER AMPLIFIER. If the frequency of the plate tank circuit of the buffer amplifier is the same as that of the oscillator driving it, the stage is a conventional type of amplifier, usually class C. If the plate tank circuit of the buffer amplifier is tuned to the second harmonic (in order to increase the frequency of the radiated signal) of the driving signal applied to the grid, the stage becomes a FREQUENCY MULTIPLIER and the output voltage has a frequency equal to twice that of the input (Figure 25-1).

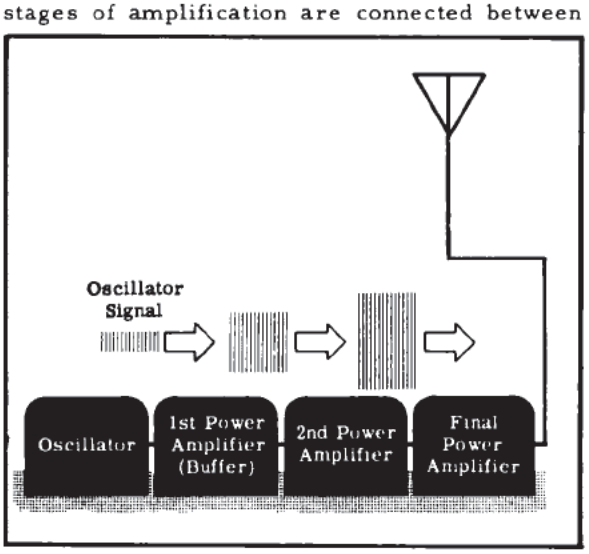


Figure 25-1 - Block diagram of transmitter.

This chapter will first review the classes of operation and types of bias used in RF amplifiers. A detailed explanation of Power Amplifiers will be presented including the tubes used for RF power amplifiers. Coupling and tuning methods will be discussed. Frequency multiplication and neutralization methods will be presented. Finally general troubleshooting information will be given.

25-1. Class of Operation

Amplifiers may be classified according to the conditions under which the tube operates—that is, according to the portion of the ac signal voltage cycle during which the plate current flows as controlled by the bias on the grid. The three classes of amplifier operation according to bias are A, B, and C, illustrated in Figure 25-2. In class A operation, the grid is biased near the midpoint of the linear portion of the plate current—grid voltage curve. The ac signal on the grid causes the grid voltage to vary above and below the bias value. The current variations are proportional to the grid voltage since the grid voltage swing does not go beyond the linear portion of the curve. Plate current flows throughout the entire ac cycle since the grid voltage does not bring the tube into cutoff. The principal characteristics of class A amplifiers are minimum distortion, low power output for a given tube (relative to class B and

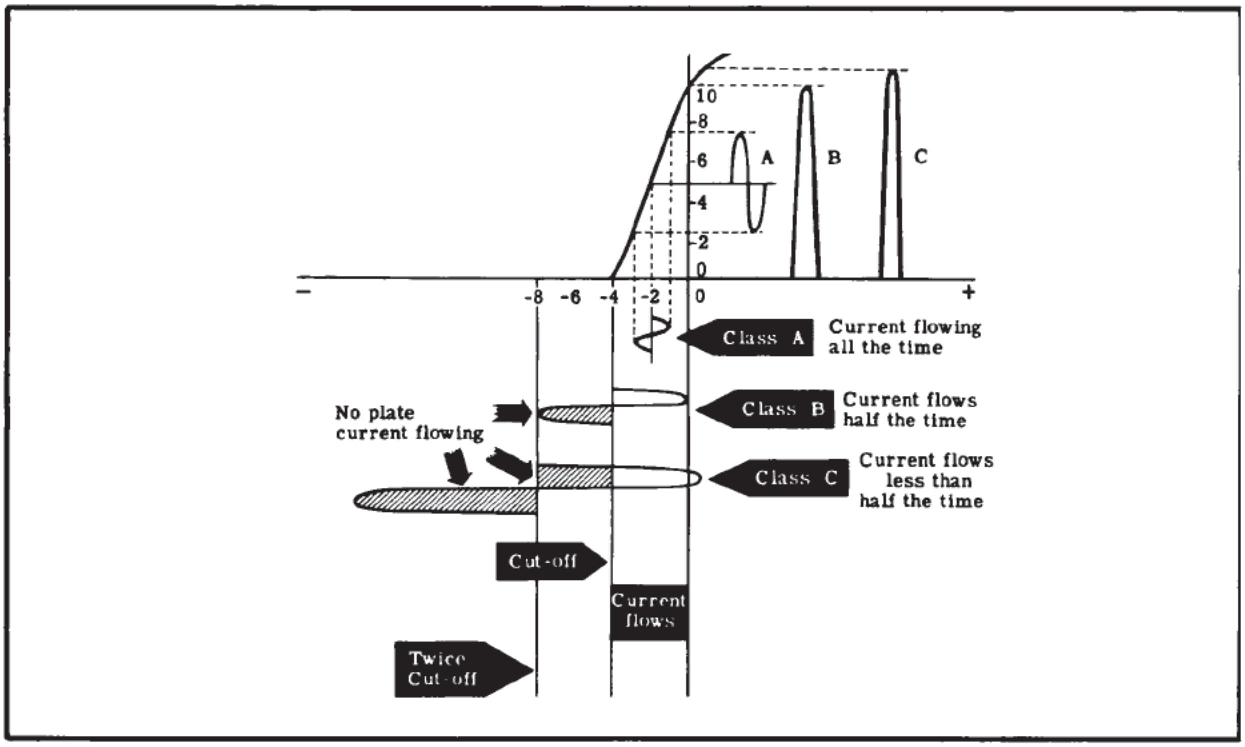


Figure 25-2 - Operation of class A-B-C amplifiers.

class C amplifiers), and relatively low plate efficiency (20-25 percent). This type of amplifier finds wide use in various audio systems where low distortion is important. In class B operation, the grid is biased at or near its cutoff value. The ac signal drives the tube into cutoff for approximately half of the cycle. Thus the tube conducts for about 180 degrees of the cycle and is cut off during the other 180 degrees of the cycle. Such amplifiers are characterized by medium power output, medium plate efficiency (40 to 60 percent), and moderate power amplification.

In class C operation—the type of operation with which you will be most concerned in your study of transmitters—the grid is biased considerably beyond cutoff. The tube remains cut off for most of each ac cycle and current flows in the tube only when the ac signal increases the grid voltage above cutoff. The plate current therefore flows in pulses as shown. This class of amplifier has a relatively high plate efficiency (70 to 80 percent) and high power output.

BIASING METHODS

Bias is defined as the average dc difference in potential existing between the control grid and cathode of a vacuum tube. The types of bias used with vacuum tubes are fixed, grid leak, cathode bias or combinations of these types. The value of the bias and the type of tube to which it is applied determines the class of operation of the amplifier circuit.

25-2. Fixed Bias

The term fixed bias describes any method of obtaining bias in which the bias remains fixed as the strength of the input signal varies. Fixed bias may be obtained from a negative power supply or from a battery. Each of these methods will keep the grid at a constant negative dc voltage which will not vary regardless of the strength of the signal input. The fixed negative bias is called C- just as the positive supply voltage is called B+. Figure 25-3 illustrates a fixed bias battery located in the grid circuit of the amplifier. The negative potential applied to the grid will limit the current flow from cathode to plate and will establish the operating point. When a signal is applied, its normal reference (zero) is changed to a reference level equal in magnitude to the value of bias voltage. The disadvantage of fixed bias is that the gain of the amplifier remains constant so that if the grid signal varies in amplitude, the output will similarly vary.

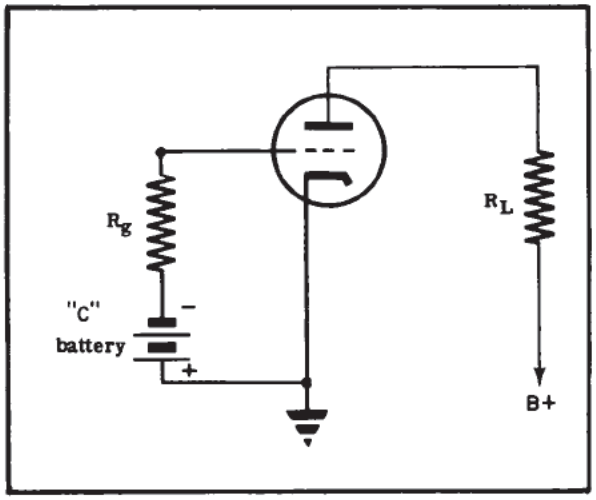


Figure 25-3 - Fixed bias.

25-3. Grid Leak Bias

A very common type of bias arrangement found in transmitters makes use of the current that flows from the cathode to the grid at the positive peaks of the input signal. This is called GRID LEAK BIAS. When the signal drives the grid positive with respect to the cathode, the grid draws current and charges up capacitor C_1 to the peak value of e_g (Figure 25-4) so that the plate of the capacitor connected to the grid becomes negative. The charge path is through the tube. Resistor R_1 provides a path for C_1 to discharge slightly between the pulses of grid current flow.

The main advantage of this type of bias is that it develops a voltage whose amplitude depends upon the strength of the input signal. If the input signal increases the grid will draw more current and the bias will become more negative. After the new value of bias has been

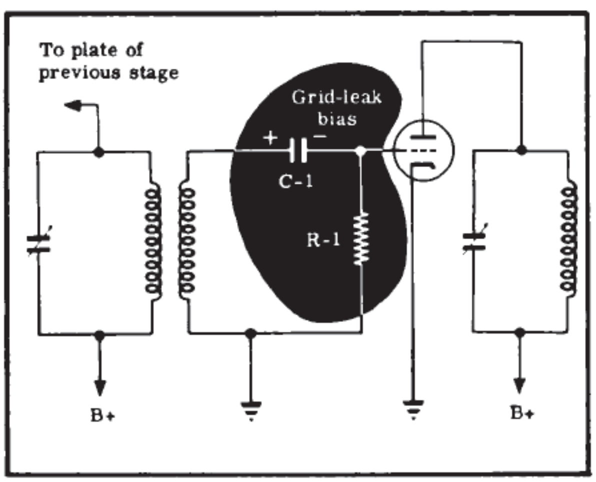


Figure 25-4 - Grid leak bias.

established, the peaks of this larger signal will not drive the grid very much more positive than a weaker signal would. Thus the peaks of the larger signal will cause about the same amount of plate current to flow as the peaks of a smaller signal. In this way, grid leak bias provides for amplitude stability.

The main disadvantage of grid leak bias is that it depends entirely upon the presence of a signal in order to develop any bias voltage, and therefore does not protect the tube when there is no signal on the grid. If the oscillator of a transmitter stopped oscillating for any reason, the grid leak arrangements in the amplifiers would not develop any bias since the grid would not, under these conditions, be driven positive. The transmitting tube would draw a very large current with zero bias and would burn out in a short time.

25-4. Combination Bias

The most common bias arrangement in transmitters is a combination of fixed bias and grid leak bias. The fixed bias is sufficient to limit the current to a low value or even to cut off in the absence of a signal. When a large enough signal is present to drive the grid positive, grid leak bias is developed which stabilizes the amplitude of the output. Thus combination bias protects the tube and stabilizes the output.

25-5. Angle of Plate Current

The angle of plate current is defined as the interval measured in degrees of the input signal during which the tube conducts. The symbol used to represent the angle of current is θ_p.

A class A amplifier is characterized by tube current for 360° of the input signal due to being biased on the linear portion of the dynamic curve. Therefore, the angle of plate current in a class A amplifier is 360 degrees. A class B amplifier is biased at or near cutoff, therefore the angle of plate current is 180 degrees because the tube is conducting for one half the time. The angle of plate current in a class C amplifier is less than 180 degrees because the tube conducts for appreciably less than half of each cycle of the applied grid signal voltage. The angle of plate current in a class C amplifier is usually 120 to 160 degrees.

Since the angle of plate current is determined by the class of operation, the class of operation determines the percentage of efficiency. The power we supply to an amplifier is always greater than the power we get out of it. Some power is used up by the tube and the rest appears as useful output in the load. The power used up by the tube equals its plate voltage times its plate current. Since the angle of plate current of a class C amplifier is less than 180 degrees, the average

plate current is less than in class A or B operation. Therefore, less power is used up by the tube and more power can get to the output. This makes the class C amplifier more efficient and therefore more desirable for use in a transmitter. The amount of plate current that is passed through the amplifier governs the power output of the amplifier stage. If the angle of plate current was decreased by increasing the bias, the average plate current is less resulting in lower power output.

Therefore, the angle of plate current will affect the output power and efficiency of an amplifier stage. A smaller angle results in greater efficiency but lower power output. An angle of 120 to 160 degrees is a compromise between these two factors.

25-6. DC Component

Class B amplifiers have characteristics in common with half wave rectifiers. The angle of plate current in both of these circuits is the same; that is, 180 degrees. For one complete cycle of class B amplifier input, there is one alternation in the output. The output of the half wave rectifier and the output from an amplifier operated class B is a train of half wave pulses. The average value of the output current for both types of circuits is:

$$I_{ave} = 0.318 \times I_{max} \qquad (16\text{-}5)$$

The peak or maximum current in equation 16-5 is the maximum value of output current.

In the class C amplifier, the plate current may also be thought of as having a dc component. It, like the current output from a class B amplifier, has an output which does not swing below the reference level. The average dc current level from a class C amplifier should be less than that of a class B amplifier. If the angle of plate current is assumed to be 120 degrees, the average plate current may be computed using the following equation:

$$I_{dc} = 0.276 \times I_{max} \qquad (25\text{-}1)$$

Where:

I_{dc} = average dc value of plate current.

If the angle of plate current is 150 degrees, the average value of plate current is:

$$I_{dc} = 0.308 \times I_{max} \qquad (25\text{-}2)$$

Derivation of the constants (0.276 and 0.308) is yond the scope of this text.

Q1. What is the average current for an amplifier which is operated class A.

25-7. Current Relations of Class C Amplifier

The operation of a class C amplifier will become clear when you analyze what happens in a tuned amplifier such as the one shown in Figure 25-5. The relationship between grid voltage, grid current and plate current is illustrated. The triode amplification factor, μ, is 20 and the plate supply voltage is 1000 volts. The cutoff bias, E_{co}, is therefore

$$E_{co} = \frac{-E_{bb}}{\mu} \cong \frac{-1000v}{20} = -50 \text{ volts} \qquad (25\text{-}3)$$

This means that the amount of bias required between the grid and cathode to cutoff the tube is equal to -50 volts. For class C operation, the operating bias, E_{co}, is normally two to five times cutoff.

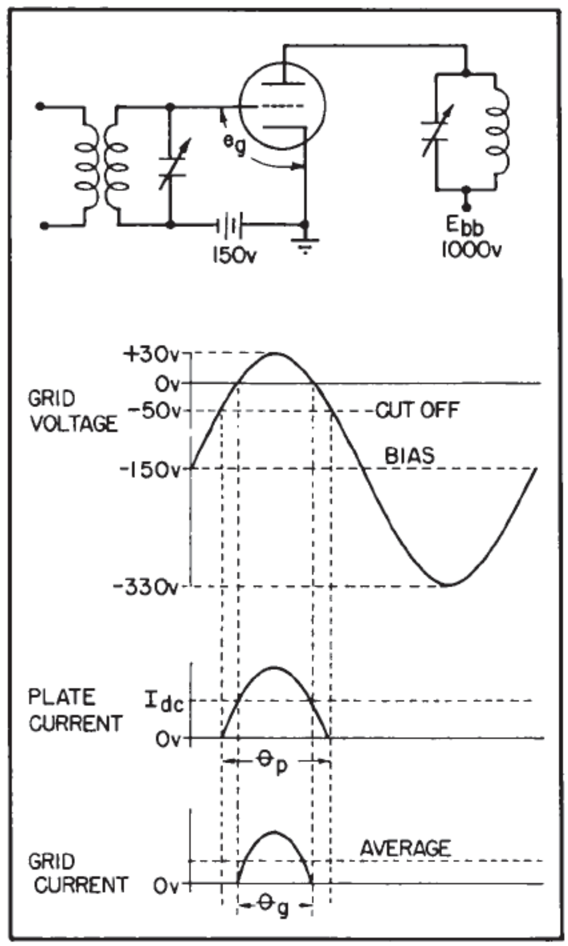


Figure 25-5 - Class C amplifier with current waveforms.

Using the value of 3 times cutoff and E_{co} of -50 volts, applied bias is -150 volts. The fixed bias battery provides this voltage for proper class C operation. The voltage actually applied to the grid electrode of the tube consists of the grid bias plus the exciting voltage. If the peak amplitude of the exciting voltage is 180 volts, when the grid end of the RF input tank circuit is positive, the peak positive grid to cathode voltage is 180-150 = +30v. When the grid end of the RF input is negative, the peak negative grid to cathode voltage is (-180) + (-150) = -330 volts.

When the grid voltage is above the cutoff value (-50v) plate current flows. When the grid becomes positive with respect to the cathode, grid current will flow as illustrated in Figure 25-5. If grid leak bias were to be utilized, this flow of grid current would produce the necessary bias for class C operation. The duration of grid current flow measured in degrees is the ANGLE OF GRID CURRENT and is labeled θ_g. If the distorted plate current waveform is broken down into its associated harmonics, it is found to consist primarily of the fundamental, second and third harmonics.

Q2. What is the angle of grid current for class B1 operation? Why?

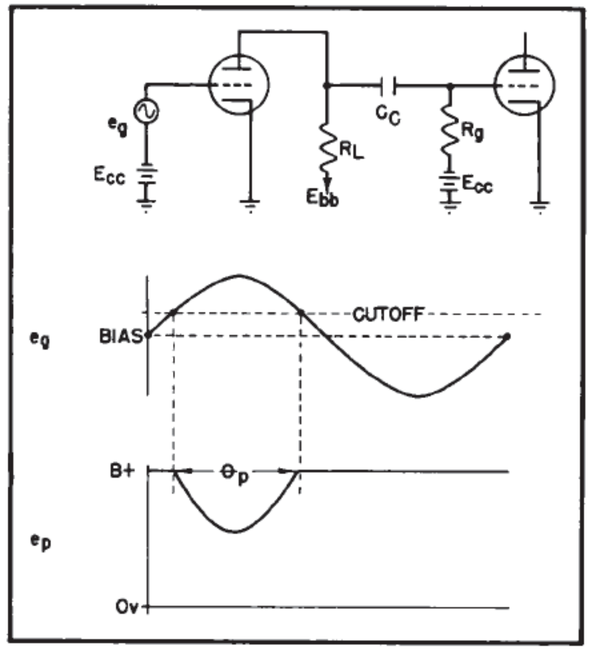


Figure 25-6 - Class C amplifier with resistive load.

AMPLIFIER LOADS

The output voltage waveforms with three different types of loads will now be compared. Each of the circuits will be operated class C. The types of loads are: RESISTIVE, INDUCTIVE and RESONANT. It will be shown that the type of load most frequently used is the resonant load.

25-8. Class C Amplifier With Resistive Load

The class C amplifier with a resistive load is shown in Figure 25-6. With the tube biased beyond cutoff, zero current flows through the tube and plate voltage equals E_{bb}. Above cutoff, tube current flows causing a voltage drop across the resistive load. This voltage drop subtracts from E_{bb} causing the plate voltage (B+) to be lower than the cutoff value. When the input signal goes below the cutoff value again, current through the tube ceases and plate voltage rises to E_{bb}. Figure 25-6 illustrates the resultant plate voltage waveshape utilizing a resistive load. This circuit arrangement is not utilized in a transmitter because the output voltage waveforms consists of pulses rather than cycles but was presented here because of its simplicity of operation.

25-9. Class C Amplifier With Inductive Load

Figure 25-7 shows an amplifier operated class C with an inductive load.

As a starting point for this analysis, assume that the tube is cut off (plate current is zero). The plate voltage at cutoff is equal to the value of the supply voltage. At this time, all of the supply voltage is felt across the tube.

Having thus established the conditions when the tube is cut off, the Thevenin's equivalent circuit will be used to analyze the circuit when the tube is brought out of cutoff. The equivalent circuit is shown in Figure 25-8. The coupling capacitor is not shown in this equivalent circuit since its reactance is negligible at the chosen operating frequency.

As soon as the signal applied to the grid becomes sufficiently positive to bring the tube out of cutoff, plate current begins to flow and the voltage felt across the inductance rises rapidly,

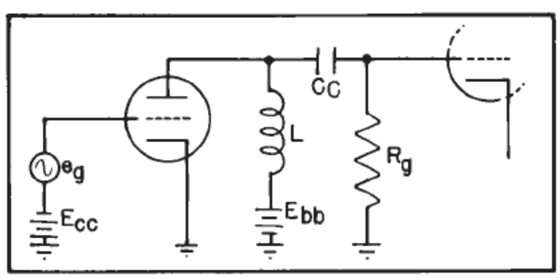


Figure 25-7 - Inductive load.

A1. The value of quiescent plate current.

A2. Zero. There is no grid current flow when an amplifier is operated class B_1.

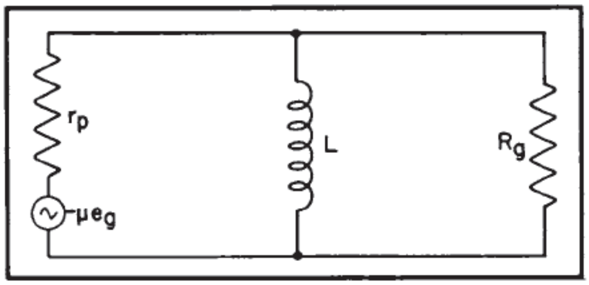


Figure 25-8 - Equivalent circuit.

causing plate voltage to drop a proportional amount. Referring to the equivalent circuit, the grid resistor of the next stage shunts the inductance.

As plate current increases, the voltage across the inductor increases making plate voltage become less positive. This negative change is coupled across R_g, providing a negative output. During this period the magnetic field about the inductor expands. Since the internal resistance of the tube is relatively low, the charging time constant (L/R) for the inductor is long.

As the grid potential becomes less positive, it rapidly approaches cutoff with a rapid decrease in plate current. This decrease in current through the inductor causes the magnetic field to collapse in an attempt to maintain current through the inductor. Once tube cutoff is reached, the discharge circuit for the inductor is through the large value of grid resistance. This constitutes a very small discharge time constant so that the current flowing through R_g is relatively large. This produces a positive plate potential (and output) which may be several times the value of the plate supply voltage. This effect, known as inductive kick, produces a very sharp spike of short duration and high amplitude.

The use of the inductive kick effect provides a means of achieving peak to peak output signals greater than the plate supply voltage. This reduces the number of stages required to achieve a driving signal of a given amplitude.

25-10. Class C Amplifier With Resonant Load

The last and most important load to be considered is the resonant load. The class C amplifier with a resonant load is shown in Figure 25-9.

The initial operation of the class C amplifier with a resonant load is similar to the operation

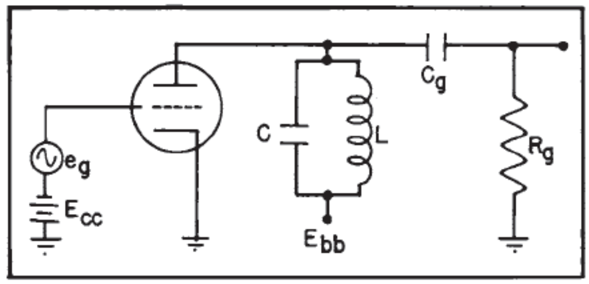


Figure 25-9 - Class C amplifier with resonant load.

of an amplifier with an inductive load. Assume that the frequency applied is equal to the resonant frequency of the tank circuit. When the tube is cutoff (zero plate current), the charge on the capacitor in the tank circuit equals zero volts and the plate voltage is equal to the supply voltage. When the grid signal reaches a value sufficient to bring the tube out of cutoff, plate current begins to flow. Since the parallel resonant circuit offers maximum impedance at the resonant frequency, a large voltage drop will be developed across the tank circuit. The capacitor in the tank circuit charges to this value of voltage. The voltage across the tuned circuit must also be equal to the same value. Maximum voltage will be dropped across the tank circuit when the current through the tube reaches its maximum value. This is shown in the waveform diagram of Figure 25-10.

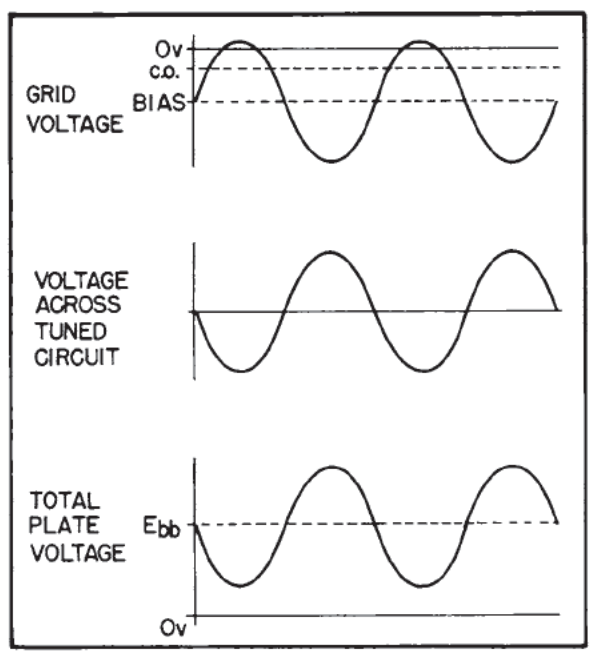


Figure 25-10 - Voltage waveforms of class C amplifier with resonant load.

Since the instantaneous plate voltage is the algebraic sum of the B+ voltage, and the ac voltage across the tuned circuit, when plate current starts flowing the voltage at the plate end of the tuned circuit goes negative and subtracts from the B+ voltage causing the plate voltage to decrease in value toward zero. When the plate current decreases, the voltage across the tank begins to decrease and the capacitor starts to discharge through the inductor. As the capacitor discharges, its discharge current sustains the field about the inductor. When the tube is sent into cutoff, plate current ceases and the field about the inductor collapses causing plate voltage to rise toward B+. The collapsing magnetic field of the inductor maintains current in the same direction through the capacitive branch of the tank circuit, charging the capacitor in the opposite direction. When the field has completely collapsed, the capacitor charge is maximum with a polarity opposite to the initial charge. When the voltage across the tuned circuit goes positive at the plate end, it adds to the B+ voltage causing the instantaneous plate voltage to be nearly double the value of E_{bb}. Once fully charged, the capacitor begins to discharge since there is no potential difference across it to sustain the charge. Its discharge path will be through the inductor. This action is known as the FLYWHEEL EFFECT and was thoroughly discussed in Chapter 12. When the grid signal again reaches a value sufficient to bring the tube out of cutoff, the tank circuit receives another burst of energy due to tube conduction at that time. This action is continuous until the grid signal is removed, at which time the oscillations (called damped oscillations) will continue until all of the energy contained in the tank is dissipated by the normal losses in the circuit. This action is extremely important, because a complete cycle of output is obtained by using only a small portion of the input cycle, and the plate voltage waveform varies above and below the B+ voltage level.

Q3. What is the relationship between the output obtained from a class C amplifier with a resistive load and an inductive load?

Q4. What is the relationship between the output obtained from a class C amplifier loaded with a resistance and one loaded with a resonant circuit?

25-11. Grid Current Loading

The grid circuit, while drawing current, causes some adverse effects that must be considered. Figure 25-11 shows the input circuit to an amplifier operated class C. For the purpose of this discussion, the feeding device will

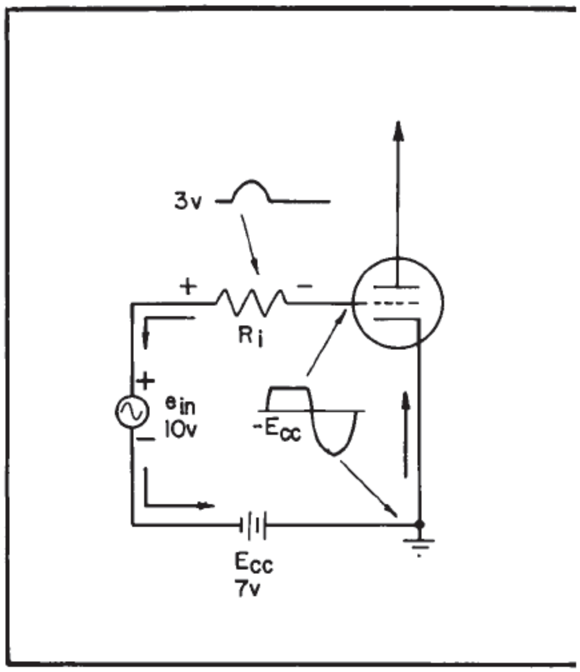


Figure 25-11 - Grid current loading.

be a generator which possesses a large value of internal resistance (R_i). The tube is biased by the negative potential E_{cc} supplied to the grid, with the cathode returned to ground. No grid current flows until the input signal e_{in} rises sufficiently to equal and effectively remove the biasing voltage E_{cc}. Any further rise of e_{in} drives the grid positive with respect to the cathode and grid current flows. The grid to cathode resistance may drop from an infinite value, when the grid is negative with respect to the cathode to a value on the order of 1000 ohms, when the grid attempts to become positive with respect to the cathode. If the internal resistance of the generator in series with the grid is 10K ohms, the drop across the 1000 ohm R_{gk} is negligible as compared to that which is developed across the 10K ohm resistor by the flow of grid current. When grid current flows through the resistor R_i, it develops an $i_g R_i$ drop of such polarity as to oppose the positive input voltage. Since the full input voltage must appear as the sum of the drop across R_i and R_{gk}, the larger R_i is with respect to R_{gk}, the nearer the voltage on the grid is limited to that of the cathode. The $i_g R_i$ drop may be considered as an automatic bias developed during the part of the input which causes grid current to flow. This particular type of distortion is known as PEAK CLIPPING with the positive peak of the input voltage signal decreased in amplitude.

A3. With the resistive load, the output signal is a partial reproduction of the input signal; whereas, the output from the inductively loaded circuit causes the plate voltage to rise much higher in value than the supply voltage.

A4. The output from a resistive load is a partial reproduction of the input signal, but the output using the resonant load yields a sine wave output.

To reduce peak clipping as much as possible, a source with a low output impedance is used.

CLASS C AMPLIFIER

25-12. Parameters of a Class C Amplifier

Since there is grid current flowing during class C operation and practical generators possess some value of internal resistance, there will be an I^2R loss in the grid circuit. The power which is dissipated in the grid circuit is called the DRIVING or EXCITING POWER. Its value is usually expressed as an average power, however, the instantaneous power dissipated at the grid will be the product of the instantaneous voltage and current. Part of the exciting or driving power will be dissipated in the form of heat at the tube grid, and part of it will be dissipated by the internal resistance of the bias battery, or, if grid leak bias is used, across the grid-leak resistor.

One other power that must be considered is the PLATE INPUT POWER. This is the power that is delivered from the power supply to the tube. To compute the value of the plate input power the following equation is used.

$$P_{dc} = E_{bb} \times I_{dc} \qquad (25\text{-}4)$$

Where:

P_{dc} = power in watts
E_{bb} = supply voltage
I_{dc} = average plate current

The smaller the angle of plate current, the more efficient is the amplifier. The angle of plate current (determined by the bias) will also affect the plate input power and the exciting power. If the plate and grid currents are permitted to flow for only a few degrees, the average plate and grid currents will be comparatively small in value. Therefore, as the angle of plate current and that of grid current decreases, the plate input power and the exciting power will also decrease.

To adequately discuss the various powers, it is best to analyze an RF amplifier using a specific high-power tube. When the values of the various powers are then found, they may be compared with the suggested values found in a transmitter tube manual. The actual circuit conditions may be established.

The circuit used for this discussion is shown in Figure 25-12. It is an RF power amplifier with a resonant load. The tube used is an 889-A high-power triode.

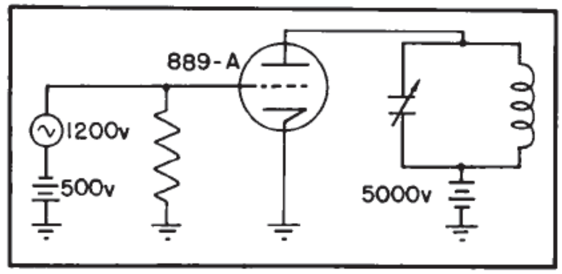


Figure 25-12 - RF power amplifier.

The 889-A finds use in commercial broadcasting and naval shore stations. The analysis of the circuit using this high power tube requires the use of a slightly different set of tube curves than used in analyzing a typical voltage amplifier. The curves used with this type of tube (889-A) are called the CONSTANT CURRENT CURVES. These curves may be found in a transmitting tube manual or they may be plotted from the plate and grid family of characteristic curves for the tube. In plotting each of these three curves, one of the tube variables is held constant, the second quantity is varied, and the effects on the third quantity are measured.

When the plate family of curves (E_b-I_b), which may also be called a constant voltage curves, are plotted; the grid voltage is held constant and the plate voltage is increased at given intervals and the plate current is measured. To obtain the first curve, the grid potential is adjusted to zero volts. Then starting at zero, the plate voltage is increased in steps, and the plate current for each of these plate voltages is recorded. From the recorded values, a curve is plotted and marked $E_c = 0$. The second, third, etc. curves are obtained by repeating the above procedure with grid potentials of -2 volts, -4 volts, etc.

If a grid family of curves (E_c-I_b), which is also a constant voltage curve, is plotted; the plate voltage is held constant and the grid voltage is increased in given intervals and the resulting plate current is then plotted for each of the grid voltage increments. If a line is drawn connecting all of these points, an E_c-I_b curve is produced.

The third family of characteristic curves is called the CONSTANT CURRENT or E_c-I_b characteristic curves and are very useful in analyzing a RF power amplifier. To construct the constant current curves, plate voltage is used as the independent variable and is plotted horizontally along the X axis. To obtain the points for the curve, the grid voltage required to produce the desired current is determined for a number of different plate voltages. The various combinations of plate voltage and grid voltage which produce the same value of current are then plotted to obtain the constant current curve.

The constant current curves for the 889-A, shown in Figure 25-13, illustrate the relations between plate voltage and grid voltage. The use of these curves makes it possible to obtain the plate voltage and plate current for various values of grid voltage. From these curves, waveforms may be plotted from which efficiency, plate dissipation, power output and other operating information can be obtained. The dashed curves found in the upper left-hand corner of the graph are grid current curves and are found by maintaining the grid current constant and varying the plate and grid voltages.

25-13. Class C Amplifier

The bias chosen for the class C amplifier (in Figure 25-12) is -500 volts which is approximately two or more times the cutoff value of the tube. If the peak amplitude of the RF input signal voltage applied to the grid of the tube is 1200 volts, the peak of the positive signal will make the grid 700 volts positive with respect to the cathode. The grid under these conditions will be capable of drawing a considerable amount of grid current. If the peak amplitude of the RF input signal is 1200 volts, the peak of the negative signal swing will make the grid 1700 volts negative with respect to the cathode which is well below the cutoff value of the tube. With the tube biased below cutoff, only a portion of the positive input voltage will drive the tube into conduction. The conduction angle will be less than 180°. These pulses of plate current will excite the resonant load in the plate circuit and the tank will go into oscillation. These surging currents give the tank circuit the so called flywheel effect, in which the tuned resonant circuit makes up the portion of the sine wave missing in the plate current pulses, and supplies a voltage of sine waveform to the load. The tube acts as a valve merely to supply the necessary power at just the right time.

The major power losses of the circuit occur in the electron tube and can be reduced by decreasing the angle of plate current. Decreasing the angle of plate current results in less plate dissipation and greater efficiency. The efficiency may also be improved by increasing the resonant impedance or raising the Q of the tank. This causes a larger voltage to be dropped across the tank and less across the tube. Less voltage across the tube (lower minimum plate voltage during conduction time) results in less tube current and hence less power lost in heat

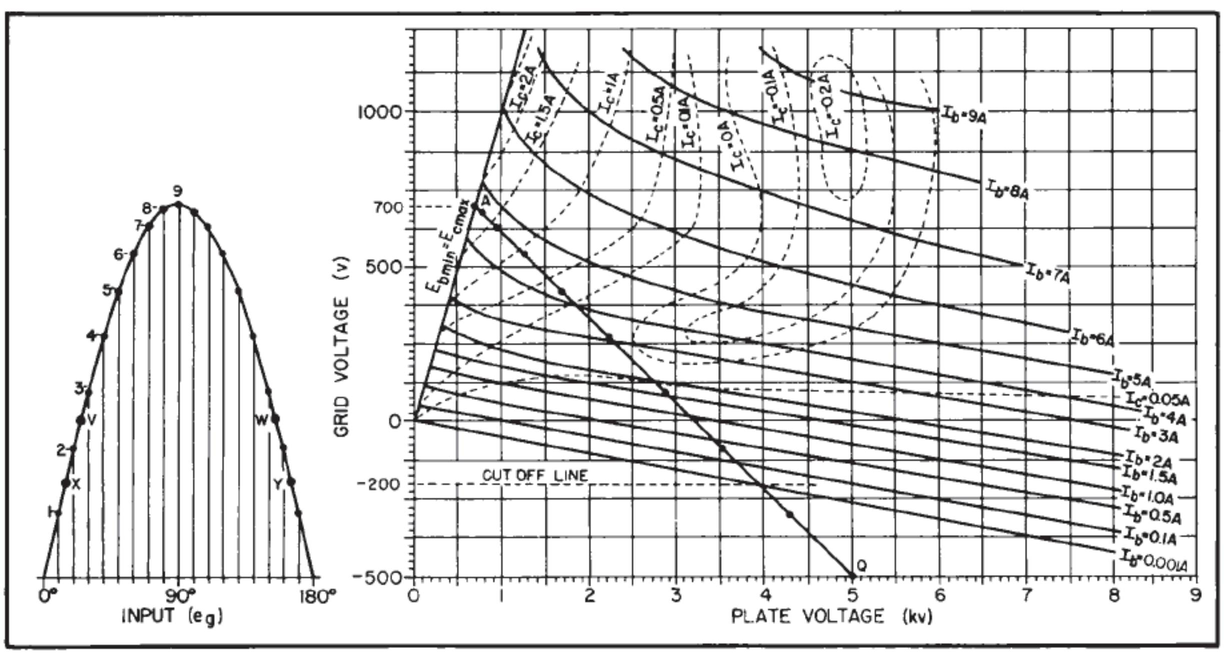


Figure 25-13 - Constant-current curves (889-A triode).

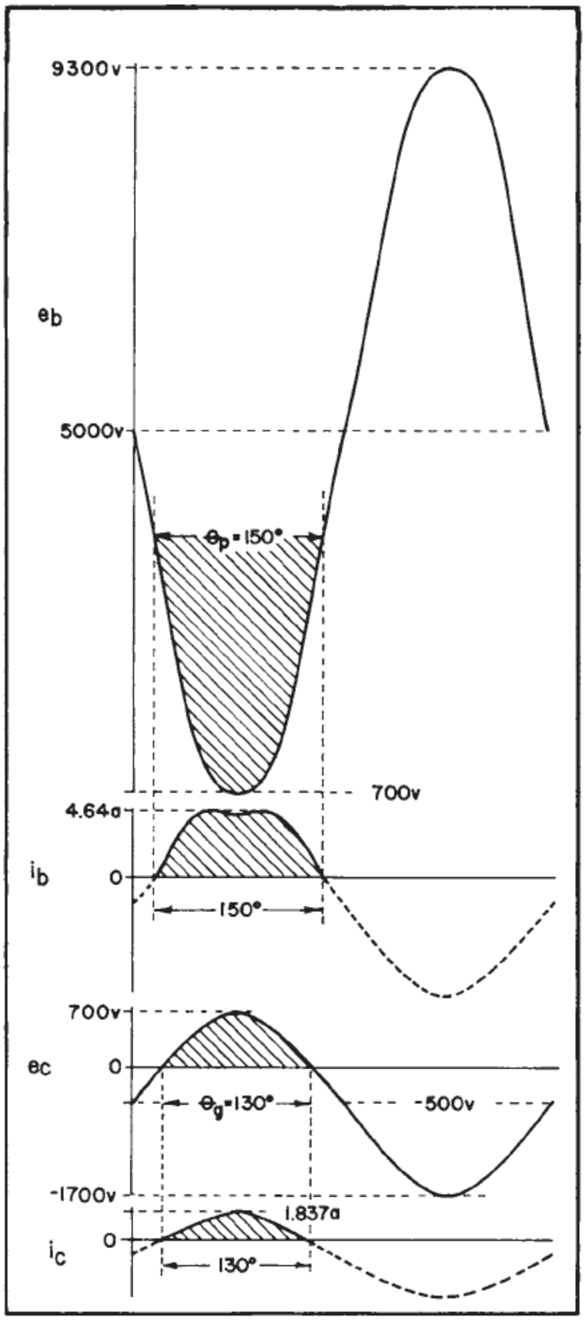


Figure 25-14 - Amplitude and phase relationships.

cuit makes up the portion of the sine wave missing in the plate current pulses, and supplies a voltage of sine waveform to the load.

Grid voltage is positive with respect to the cathode for a short period of time during each cycle. For optimum conditions the minimum value of plate voltage (e_b min), should be equal to the maximum positive value of grid driving voltage (e_c max). In other words e_c max = e_b min. The maximum positive grid voltage (e_c max) should never be allowed to exceed the minimum plate voltage (e_b min). Otherwise, plate current would decrease and grid current would become excessive, resulting in a reduction in output power and excessive grid losses. Notice that there is a dip in the peak of the plate current waveform in Figure 25-14. This dip is due to the flow of grid current and is caused by large grid drive. When the voltage applied to the grid reaches its maximum positive value of +700 volts the plate voltage is equal to +700 volts. The plate and control grid are at the same positive potential with reference to ground. The control grid, being physically closer to the cathode will attract a greater number of electrons causing a sharp increase in grid current and a dip in plate current.

Since the bias voltage is -500 volts, the input grid signal will be superimposed on that level. This means that the signal voltage will have to rise above 500 volts before the grid begins to draw current. On the negative swing of the input signal, the signal peak voltage adds to the bias voltage to make the grid -1700 volts with reference to the cathode. The grid current value is found from the curves in Figure 25-13. The peak value of the grid current is 1.837 amperes. The angle of grid current is also shown on the grid current and grid voltage waveforms.

25-15. Power and Efficiency

To compute the average values of power input, power output and grid driving power, the average values of plate current, grid current and grid voltage must be known. The values of current given thus far have all been peak values. As was pointed out in the chapter dealing with rectifier circuits, the average value of a pulse is determined through the use of calculus, a branch of mathematics that is used in advanced electronics. Due to the similarity between the plate current pulse of a class C amplifier and that of a half wave rectifier (only in so far as they are both pulsating dc), the average value of plate current could also be found by the use of calculus. However, since this is a very complicated process, a substitute method will be employed which will render an average value sufficiently accurate for this application. The average value of a current pulse for various conduction angles may be found through the use of Table 25-2. This table shows the value of the conversion factor for several angles of plate current that must be divided into the peak current to find the average.

The peak value of the current pulse of a class C amplifier found from Figure 25-13 was 4.64

θ_p	Average Plate Current
180	i_b max/3.14
160	i_b max/3.50
150	i_b max/3.75
140	i_b max/4.00
130	i_b max/4.25
120	i_b max/4.60
110	i_b max/5.00
100	i_b max/5.50
90	i_b max/6.10

Table 25-2 - Conversion factors from peak to average.

θ_p	I_b
90	I_{avg} x 1.88
100	I_{avg} x 1.85
110	I_{avg} x 1.83
120	I_{avg} x 1.79
130	I_{avg} x 1.76
140	I_{avg} x 1.72
150	I_{avg} x 1.69
160	I_{avg} x 1.65
170	I_{avg} x 1.61
180	I_{avg} x 1.57

Table 25-3 - Conversion factors from average to peak.

amperes. With a conduction angle of 150 degrees, the average value of this current pulse is:

$$\begin{aligned}\text{Average} &= i_b \text{ max}/3.75\\ &= 4.64/3.75\\ &= 1.24 \text{ amperes}\end{aligned}$$

The pulse of tube current of a class C power amplifier is a periodic nonsinusoidal waveform and as such, it is subject to Fourier analysis by which it may be resolved into a sine wave of the same frequency as the grid driving signal and many harmonics of that frequency. This information is included because the power output at the fundamental frequency, because of harmonic content, is less than the total power output. All of the harmonics combined with the fundamental constitute the total power output. The power output of the circuit at the fundamental frequency is given by the equation:

$$P_o = \frac{(E_{bb} - e_b \text{ min})\, I_b}{2} \qquad (25\text{-}5)$$

Where:

P_o = power output in watts

$(E_{bb} - e_b \text{ min})$ = peak plate voltage swing.

I_b = crest or peak value of the fundamental ac component of plate current.

2 = conversion from peak power to average power.

To determine the crest or peak value of the fundamental ac component of plate current when the average value of the current pulse is known, the appropriate conversion factors from Table 25-3 are used.

If the average value of the current pulse is 1.24 amperes, the peak value of the fundamental in amperes may be found with the aid of Table 25-3. With a conduction angle of 150 degrees, the peak plate current is:

$$\begin{aligned}\text{Peak} &= \text{average current} \times 1.69\\ \text{Peak} &= 1.24 \times 1.69\\ \text{Peak} &= 2.1 \text{ amps}\end{aligned}$$

Notice that this value of peak plate current is less than that found from the constant current curves of Figure 25-13. The reason being that this peak value is of the fundamental frequency only and not the harmonic frequencies.

Substituting values into equation 25-5.

$$P_o = \frac{(5000-700)\ 2.1}{2}$$

$$P_o = 4515 \text{ watts}$$

The plate input power may be computed using the formula:

$$P_{input} = E_{bb} \times I_{dc} \qquad (25\text{-}6)$$

Where:

I_{dc} = Average dc plate current

E_{bb} = Plate supply voltage

Substituting values

$$\begin{aligned}P_{input} &= 5 \times 10^3 \times 1.24 \text{ amps}\\ &= 6200 \text{ watts.}\end{aligned}$$

The value of plate power dissipation is the difference between the plate input power and the output power.

$$P_{dissipated} = P_{input} - P_{output} \quad (25\text{-}7)$$
$$= 6200 - 4515$$
$$= 1685 \text{ watts}$$

This means that with a dc input power of 6200 watts, 1685 watts is dissipated in the tube and circuit components in the form of heat and 4515 watts of useful power is delivered to the load.

The plate circuit efficiency may be expressed as

$$\text{EFFICIENCY} = \frac{P_{ac}}{P_{dc}} \text{ or } \frac{P_{out}}{P_{in}} \times 100 \quad (25\text{-}8)$$

$$\text{EFFICIENCY} = \frac{4515}{6200}$$

$$\text{EFFICIENCY} = 72.8\%$$

This formula illustrates that the efficiency of a class C power amplifier may be improved by decreasing the dc input power to the stage. This may be accomplished by decreasing the angle of conduction. A smaller angle of conduction results in less heat dissipation and greater circuit efficiency but less power output. The compromise between maximum efficiency and maximum output power is achieved by operating the stage with θ_p usually between 120° to 160°.

The instantaneous value of grid driving power is the product of the instantaneous values of grid voltage and grid current. To determine the actual value of the grid driving power over the complete cycle, the following equation is used:

$$P_c \cong e_c \times i_c$$

Where:

P_c = instantaneous value of grid driving power
e_c = instantaneous value of grid voltage
i_c = instantaneous value of grid current

Substituting value:

$$P_c \cong 1200 \text{ volts} \times 242 \times 10^{-3}$$
$$P_c \cong 290.4 \text{ watts}$$

The average value of grid current may also be found through the use of appropriate conversion table (not given). It should be noted that this value of grid driving power is not the maximum possible value of driving power that could be applied to the tube. It merely indicates that with 290.4 watts of driving power, there will be 6200 watts of input plate power and 4515 watts of output power. Higher values of driving power could be applied to the tube resulting in higher output power.

Q5. What happens to the power output of the fundamental frequency if the angle of plate current is decreased? Explain.

Q6. What controls the angles of grid and plate currents?

Q7. What happens to the efficiency of a class C amplifier if the peak grid swing is reduced and the bias voltage remains constant?

25-16. RF Amplifier Tubes

The choice of the type of electron tube to be used in an RF power amplifier is determined by the requirements of the circuit in which it is to be used. Tetrodes and beam power tubes enjoy wide popularity as RF power amplifiers because of their ability to pass large values of current, and have large plate voltage handling capacity.

The value of the peak plate current passed through a tube is governed by the ability of the electron emitter to supply free electrons. Ordinarily, the maximum permissible value of peak plate current is approximately equal to four times the sum of the dc plate and grid currents flowing under normal class C operating conditions. There are emitters in some types of transmitting tubes which require several amperes for emission or heating. Materials used for these elements must be carefully chosen so that they will produce the desired emission or heating with minimum power loss and maximum life expectancy.

In the 889-A medium power triode previously discussed, it was seen that the power dissipated by the plate of the tube was equal to 1685 watts. This high value of power is effectively lost in the form of heat. This quantity of heat would eventually destroy the tube unless the excessive heat were removed. Two common methods of removing this excessive heat is by the use of forced air cooling and by the use of liquid cooling.

The 889-A high power triode is cooled by water which is forced through the tube's enclosing jacket at a rate of three to six gallons per minute. The jacket temperature of the tube is never permitted to rise above 70 degrees centigrade. The cooling methods will extend the power handling capacity of a given tube and substantially increase the life expectancy of the tube.

The physical size of the tube used for a high power RF amplifier is directly related to its power handling capability. If the physical size of the tube is large, it may be expected that the tube will possess a high power handling capacity. For example, the 889-A medium power triode is approximately 11 inches high and three and a half inches in diameter.

In high power transmitting tubes, directly heated cathodes are preferred. The reason for this is that directly heated cathodes constructed from thoriated tungsten have the ability to withstand higher differences in potential than the normal barium or strontium coated indirectly heated cathodes. In the manufacture of the thoriated tungsten emitter, the thorium and the tungsten are mixed in quantities that render a wire in which the thorium and tungsten are homogeneous. When the completed filament wire is "flashed," the thorium migrates to the surface of the wire where it is very strongly bound to the tungsten atoms. Such is not the case with either the barium or strontium coated indirectly heated cathode. The barium or strontium is simply deposited on the nickel cathode with little or no chemical action taking place. The barium or strontium coating is not fused to the nickel as well as is the thorium to the tungsten. When a few thousand volts difference in potential is placed between the plate and the barium or strontium coated cathode, the electrical stress on the coating may be sufficiently high to physically pull the coating from the cathode surface. With high plate voltages, the life expectancy of these tubes is very short. In comparison, when the same amount of voltage difference is applied between the plate and a thoriated tungsten directly-heated cathode there will be no diffusion of the thorium from the tungsten. Therefore, directly heated cathode power tubes possess a longer life expectancy than their indirectly heated counterpart operated under the same conditions.

The voltage applied to the directly heated cathode may be either ac or dc. The ac fluctuations produce no significant affects upon the output signal. The amplitude of the ac voltage applied to the heater is in the range of about 11 volts. In comparison, it is not unusual to find a plate swing in a power amplifier of several thousand volts. Thus, the interference caused by the 11 volts is negligible.

Figure 25-15 shows a typical cathode circuit of a transmitting tube utilizing a directly heated cathode. The filament voltage is supplied by a

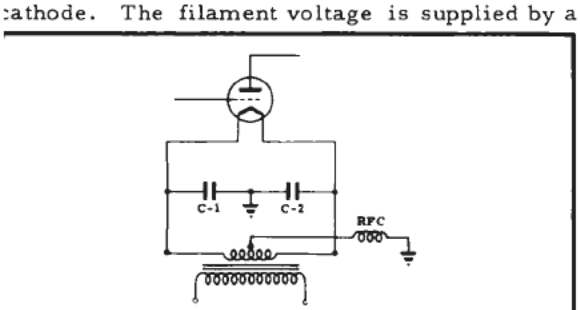


Figure 25-15 - Typical cathode circuit of transmitting tube.

transformer which is center tapped to insure symmetrical conduction of current through the tube. The untapped filament transformer shown in Figure 25-16 illustrates the non-symmetrical conduction in the cathode circuit when the filament transformer is not center tapped. As the

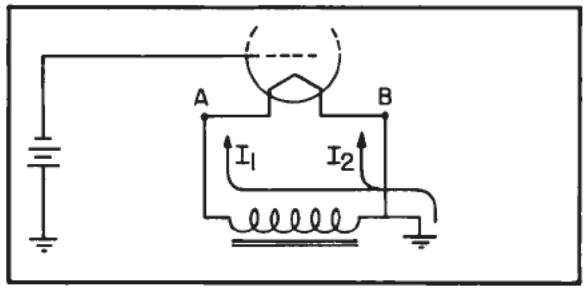


Figure 25-16 - Untapped filament secondary.

potential applied to the filament transformer varies an alternating current flows through the filament wire causing thermionic emission. There are two paths for cathode current between ground and the filament, as represented by I_1 and I_2. One of the paths is from ground through point B and through the tube. The other current path is from ground through the secondary winding of the transformer, through point A and through the tube. Even though the secondary of the transformer possesses a small value of resistance, the high value of tube current required in high power transmitters will cause an appreciable voltage drop across the filament transformer secondary winding. This means that the emission of electrons in the left hand side of the filament will be less than the emission from the right hand side since point A is positive with respect to point B. To avoid this inequality in emission, the transformer is center tapped so that the dc tube current that flows through the RFC divides in going through the filament transformer winding. Since the dc current divides, both ends of the filament are at the same dc potential, resulting in equal amounts of current to flow in each half of the filament wire. Center tapping the secondary winding increases the life expectancy of the tube.

The center tap of the transformer in Figure 25-15 is connected to ground through the RFC to keep the RF current from flowing in the transformer winding. The path of RF current and the filament is through the low reactance of C_1 and C_2.

Q8. What is the purpose of using thoriated tungsten rather than oxide coated cathodes?

A5. The average power output will decrease because the average value of plate current will decrease.

A6. The value of the bias voltage, and the peak input signal.

A7. The efficiency increases, but the power output decreases.

A8. At high plate voltages, the life expectancy of the tube is greater.

COUPLING

A single stage of power amplification normally is not sufficient for radio transmitters. To obtain the necessary gain, several stages must often be connected together. The output of one stage then becomes the input of the next throughout the series of stages and this arrangement is called a cascade amplifier. A cascaded RF amplifier is designated according to the method used to couple one amplifier stage to the next. There are a number of methods each having certain advantages and disadvantages and the choice for a particular application depends on the needs of the circuit. The basic methods are (1) resistance capacitance coupling; (2) impedance coupling, (3) transformer coupling and (4) link coupling.

25-17. Resistance Capacitance Coupling

One of the most widely used methods of connecting amplifier stages is RC coupling. Amplifiers coupled in this manner are relatively inexpensive, are relatively free from undesirable induced currents from ac heater leads and are especially suitable for use with pentodes and high mu triodes.

A typical RC coupled amplifier circuit is shown in Figure 25-17 together with the names of the various circuit elements.

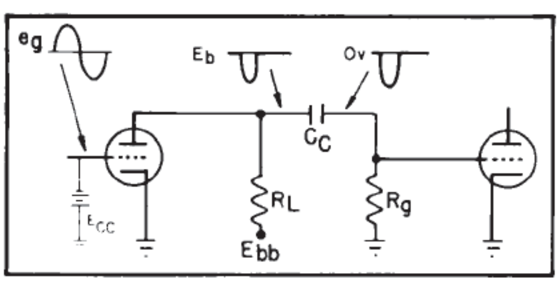


Figure 25-17 - Class C amplifier utilizing RC coupling.

In order that the output voltage may be large, the load resistor should have as high a value as practicable. However, the higher this value becomes, the greater is the voltage drop across it and the lower is the voltage remaining between the plate and cathode of the tube. Thus, there is a practical limit to the size of the plate load resistor if the plate is to be supplied with its rated voltage.

Even though the RC coupled amplifier is well suited for the job of amplifying a wide range of frequencies, there are still causes for a drop in gain at high frequencies. The reduction in gain is caused by the shunting effect across load resistor (R_L) by the plate-to-cathode interelectrode capacitance of the tube from which the signal is taken, the grid-to-cathode capacitance of the tube to which the signal is brought and the stray capacitance and inductance between the wiring and the chassis. The effect of this stray capacitance is to shunt the plate and grid resistors. At low and medium frequencies the reactance of these small capacitances is high and therefore it does not disturb the operation of the circuit. At high frequencies, the reactance decreases and effectively decreases the impedance between grid and ground. The signal appearing between grid and ground decreases as the impedance between grid and ground decreases.

RC coupling is not used with a power amplifier because of the excessive I^2R loss caused by the high value of plate load resistance. Also, without a tuned tank circuit a sinusoidal output waveform cannot be produced with class C operation.

Q9. What causes the gain of a RC coupled amplifier to decrease at low and high frequencies?

25-18. Untuned Impedance Coupling

Impedance or inductive coupling is obtained by replacing the load resistor (RL), of a normal RC coupled amplifier with an inductance (L) as shown in Figure 25-18.

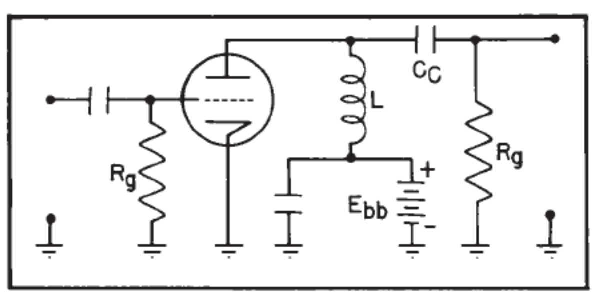


Figure 25-18 - Impedance coupled amplifier.

To obtain as much amplification as possible, particularly at low frequencies, the inductance is made as large as practicable. To avoid undesirable magnetic coupling, a closed shell type of inductor is used. Because of the low dc resistance of the inductor, less dc voltage appears across it. Thus, the tube can operate at higher plate voltage with the same power supply providing the maximum rated plate voltage is not exceeded.

The degree of amplification is not as uniform as it is with RC coupling because the load impedance Z_L varies with the frequency, this is

$$Z_L = R + j\, 2\pi f L$$

Since the output voltage appears across Z_L, the voltage gain increases with the frequency up to the point where the shunting capacitance limits it. The shunting capacitance includes not only the interelectrode capacitance and the distributed capacitance found in RC coupled amplifiers, but also the distributed capacitance associated with the turns of the inductor. The distributed capacitance between the turns of the coil greatly increases the plate to ground capacitance and plays a major part in limiting the use of this type of coupling at higher frequencies.

Generally, impedance coupling is not used with a high power RF amplifier operated class C because the inductor does not faithfully produce a sine wave which is the action that takes place when a resonant circuit is used in the plate circuit. This type of coupling is also known as untuned impedance coupling.

25-19. Tuned Impedance Coupling

Tuned impedance coupled RF amplifiers find wide application in transmitters because of their amplification, selectivity and ability to create a sine wave output. Figure 25-19 illustrates two types of tuned RF amplifiers and their performance is virtually the same. The difference between them is in the manner in which energy is supplied to the tank circuits.

Figure 25-19A shows a typical circuit for triode tube using series plate feed; that is, the dc component of plate current flows through the dc supply and the tuned circuit in series. A resonant circuit is used as the load and is tuned to the frequency of the signal impressed on the grid of the tube. The resonant circuit presents a high impedance at this frequency, whereas its impedance is lower at all other frequencies, decreasing rapidly as the frequency shifts from the resonant value. The high value of plate voltage across the tuning capacitor presents a danger to the operator making tuning adjustments which limits the usefulness of this circuit in high power transmitters.

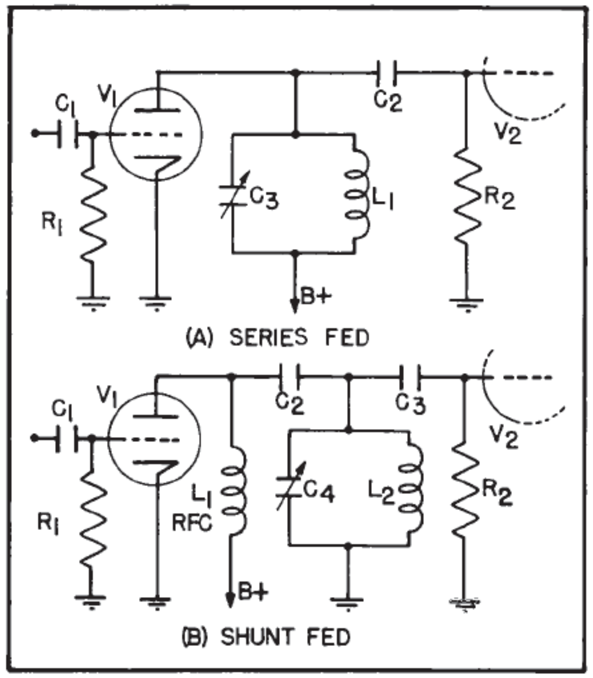


Figure 25-19 - Series and shunt fed RF amplifiers.

A more commonly used form of class C amplifier circuit is shown in Figure 25-19B. It differs from the series feed circuit in that the plate tank circuit is connected in a position where it shunts the plate load (RF choke). In this position, the full plate voltage signal is still developed across the tank circuit without passing the high dc component of plate current through it. The amount of plate current may still be changed by varying the impedance of the tank circuit. This is true because the tank circuit is in parallel with the RF choke and thereby has an effect on the total impedance offered to the operating frequency. The danger of the high value of plate voltage present during adjustment is eliminated when using the shunt fed amplifier. However, there still is a hazard involved because of the RF frequencies present. RF burns can be just as lethal as the high voltage. The shunt fed amplifier does not require the amount of insulation that the series fed amplifier requires.

25-20. Untuned Secondary-Transformer Coupling

Inductive or transformer coupling is a very efficient means of transferring energy from one stage to the next. This method of coupling inherently discriminates against harmonics of the operating frequency and hence is particularly suitable as a coupling device.

A9. Fall off at low frequencies is caused by the high reactance of the coupling capacitor. High frequency attenuation results from the shunting effect produced by plate to cathode and grid to cathode capacitance of the amplifier tubes.

Figure 25-20 illustrates an untuned transformer coupled amplifier. Two coils labelled L_1 and L_2 constitute an air core transformer. Air core transformers are used exclusively in transmitters due to the operating frequencies involved. At high frequencies the chief difficulty with iron cores transformers is increased by eddy current losses. Eddy current losses in iron cores are proportional to the square of the frequency and to the square of thickness of the laminations in the core which becomes very great at high frequencies. These eddy currents in the core also induce currents of their own which act as a shield and prevent the main magnetic flux from penetrating very deeply into the laminations or particles of iron in the core and thus causes the effective permeability of the iron to decrease with increased frequency.

In both of the stages shown in Figure 25-20, grid leak bias is developed in the usual fashion. The plate tank is tuned to the operating frequency. When a positive voltage of sufficient amplitude is applied to the grid of the tube, plate current will commence to flow. The output of the amplifier taken from the plate of the tube will be a sine wave at the plate tank resonant frequency. A result of the flywheel action of the plate tank circuit.

A characteristic of a parallel resonant circuit is that it offers maximum impedance to the resonant frequency causing plate current to be minimum. Therefore, the line current is minimum, but the tank current, the circulating current within the tank, is maximum. The maximum circulating current passing through the inductor L_1 will cause a magnetic field to fluctuate about it. Since the inductor L_2 is in close proximity, it also will be under the influence of the same field. Therefore, a voltage will be induced into the inductor L_2, through basic transformer action. The amount of coupling between L_1 and L_2 may be varied by physically moving the secondary of the transformer closer to or farther away from the primary depending upon the amount of drive required to the next stage.

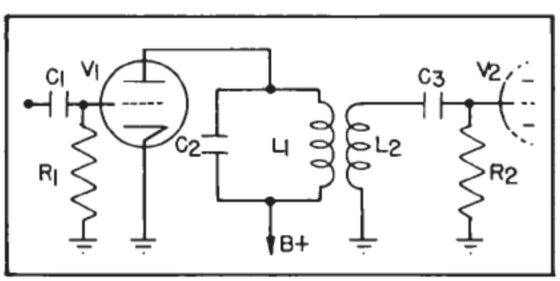


Figure 25-20 - Inductively coupled amplifier.

25-21. Tuned Secondary-Transformer Coupling

If capacitors are placed across the primary and secondary windings of the transformer in a transformer-coupled network, a double-tuned transformer-coupling system is obtained, Figure 25-21. Coil L_1 is the primary and coil L_2 is the secondary of the transformer. C_2 tunes L_1 to resonance at the signal frequency. A large signal voltage is produced across the high impedance of the parallel-resonant circuit formed by L_1 and C_2. The large circulating tank current in the primary of the transformer creates a magnetic field which induces a voltage in the secondary winding, (L_2).

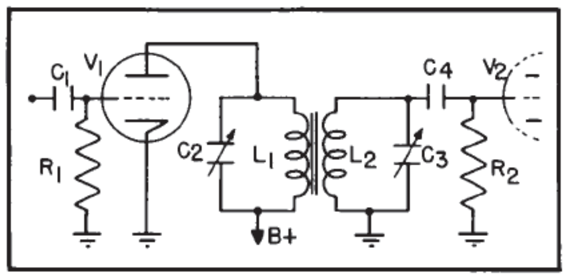


Figure 25-21 - Double-tuned transformer coupling.

The voltage coupled to the secondary circuit, L_2C_3, by induction is considered to be in series with the components of this circuit. When the secondary circuit is tuned to resonance, a large current flows which is in phase with the induced voltage. Both this current and the induced voltage are 180° out of phase with the primary voltage as in any transformer. The large secondary current produces a large reactive voltage drop across C_3, which is applied to the grid of V_2. This reactive voltage lags the secondary current by 90°.

25-22. Link Coupling

Link coupling is a special form of inductive coupling. It requires the use of two tuned circuits, one in the plate circuit of the driver tube and the other in the grid circuit of the amplifier. A low impedance RF transmission line having a coil of one or two turns at each end is used to couple the plate and grid tank circuits. The coupling links or loops are coupled to each tuned circuit at its cold end (point of zero RF potential). Circuits which are cold near one

end are called unbalanced circuits. Link coupling systems normally are used where the two stages to be coupled are separated by a considerable distance. One side of the link is grounded in cases where harmonic elimination is important or where capacitive coupling between stages must be eliminated, Figure 25-22.

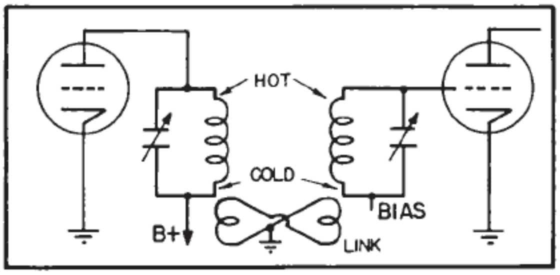


Figure 25-22 - Link coupling.

Link coupling is a very versatile interstage coupling system. It is used in transmitters when the equipment is sufficiently large to permit the coupled coils to be so positioned that there is no stray capacitive coupling between them. Link circuits are designed to have low impedance so that RF power losses are low. Coupling between the links and their associated tuned circuits can be varied without complex mechanical problems. These adjustments provide a means of obtaining very low coupling between stages.

Q10. What are the advantages and disadvantages of interstage transformers?

25-23. Amount of Coupling

When two coils are so placed in relation to each other, that the magnetic lines of force produced by and encircling one coil, link the turns of the other, the coils are said to be inductively coupled. If an ac voltage is applied to one coil, an ac voltage will be induced in the second coil. This effect of linking two inductors is called mutual inductance and the amount of mutual inductance between two coils depends upon their size and shape, their relative position and the magnetic permeability of the medium between the two coils. The extent to which two inductors are coupled is expressed by a coefficient of coupling.

From transformer theory it is known that the primary voltage will induce a voltage in the secondary. It should also be understood that when a load is connected across the secondary, and secondary current flows, the presence of the loaded secondary will cause a counter emf to be induced in the primary circuit of such a polarity as to reduce primary circuit current. This action causes the primary to act as if the resistance of the secondary circuit were reflected back into the primary. The value of this APPARENT reflected resistance is partially dependent on the amount of coupling between the two circuits. When the two circuits are physically separated, or positioned, so that very little coupling exists, the secondary circuit has little effect on the primary and the reflected resistance appears small. When the circuits are positioned in such a manner that a large amount of coupling exists, the secondary circuit has a great effect on the primary and the reflected resistance appears large.

To demonstrate the effect of coupling variations, the two coils of the transformer are first moved far enough apart so that very little coupling takes place (low value coefficient of coupling, k). This condition is called LOOSE or UNDERCOUPLING. With loose coupling there is very little transfer of energy between the primary and secondary circuits resulting in insufficient grid drive. The value of induced secondary voltage is small because the value of mutual inductance (M) is small. The low value of M also causes the value of reflected resistance to be small. The small value of reflected resistance has very little effect on the operation of the primary circuit. Due to the small amount of interaction between the primary and secondary, the two circuits behave essentially as if they were separate tuned circuits.

As the coefficient of coupling is increased (the circuits are continually moved closer together) the value of mutual inductance increases. This continually increases the amount of voltage induced in the secondary winding. The value of the reflected impedance is also increasing and due to its effect on the primary current, the response curve becomes wider (bandwidth increases).

As the coefficient of coupling is continually increased, the reflected impedance, which is resistive at the resonant frequency, continues to increase until eventually a value of coupling is reached where the reflected resistance equals the equivalent resistance of the primary. This condition is called CRITICAL COUPLING (kc) and since at this point the reflected resistance matches the primary resistance, there is a maximum transfer of energy between the circuits. Thus, the induced secondary voltage is at its maximum value as shown by the response curve for critical coupling in Figure 25-23. Notice that the critical coupling curve is flattened slightly on top.

When the coefficient of coupling is increased beyond the value of kc, the value of primary current at resonance begins to decrease. This is caused by the continued increase of the re-

A10. Advantages are (1) step up ratio of transformer permits the amplifier voltage gain to exceed the tube mu (2) a lower plate supply voltage is required. Disadvantages are (1) cost (2) poor frequency response (3) stray ac fields induce undesirable stray voltages in the transformer.

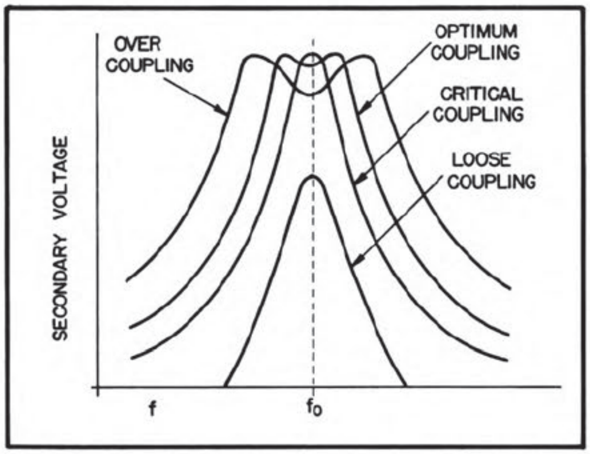


Figure 25-23 - Response curves for various amounts of coupling.

flected impedance (which at resonance is resistive) and is indicated by the dip in the response curve at f_o of the OPTIMUM COUPLING curve in Figure 25-23. A continued increase in the coefficient of coupling will result in further reduction of the gain at resonance and a further increase in the distance between the two resonant peaks (above and below f_o). The circuits are now said to be OVERCOUPLED and the overall response of the stage will take on the appearance of the double humped curve in Figure 25-23. In most equipments, the amount of coupling is not variable, but is a fixed value established during manufacture. The equipment is designed so that the amount of coupling is satisfactory over the operating frequency range.

25-24. Tuning Methods

If a class C amplifier is to operate efficiently, the plate tank circuit must resonate at the same frequency as the grid signal. If the tuning capacitor of the plate tank circuit is variable, illustrated in Figure 25-24A, the plate circuit will be either on or off resonance depending upon the setting of the variable capacitor. Adjusting the variable capacitor to make the plate tank circuit resonant to the grid signal is called TUNING. When a transmitter is detuned, a weak signal will be radiated and receivers tuned to the transmitter frequency may not pick up

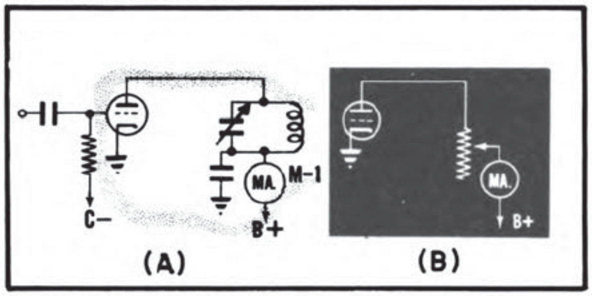


Figure 25-24 - Tuned class C amplifier circuit.

the signal. When a transmitter is tuned to a given frequency, all the tank circuits in the transmitter are tuned to resonate at this given frequency. The transmitter then radiates a stable signal at maximum efficiency and maximum power output.

A tank circuit in series with the plate of a class C amplifier can be compared to a rheostat (Figure 25-24B) in series with the plate. When the plate tank circuit is completely detuned, it acts just as if there were no resistance in the plate circuit because the impedance of the tank is minimum. As a result, plate voltage will always be equal to B+ and the pulses of plate current (when grid is driven above cutoff) will be large. The dc meter (M_1) which measures the average of the current pulses will therefore read high.

As the tuning capacitor (Figure 25-24A) is varied so that the resonant frequency of the tank circuit comes closer to the grid signal frequency, the impedance of the plate circuit increases. Now a signal voltage appears across this impedance. Just as in an ordinary amplifier, when the grid signal is positive, the plate voltage drops because of the voltage drop across the plate load resistor. Since the plate voltage is now lower than before, (lower than B+) during the time the grid is driven above cutoff, the pulses of plate current will be lower in amplitude, and therefore their average value will be less. When the plate tank is tuned to the grid signal, the plate impedance is at its highest point and therefore the voltage drop across the impedance is at its highest point. As a result, the plate voltage (the difference between B+ and the load voltage) is at its lowest point. Since the plate voltage is at its lowest point (during the time the grid is above cutoff), the plate current pulses and therefore the average plate current will be at their lowest point. This sequence of tuning is illustrated in Figure 25-25.

A minimum dc plate current reading is therefore an indication that the plate tank is tuned to the grid signal frequency. When a plate tuned circuit is tuned for a minimum reading on the

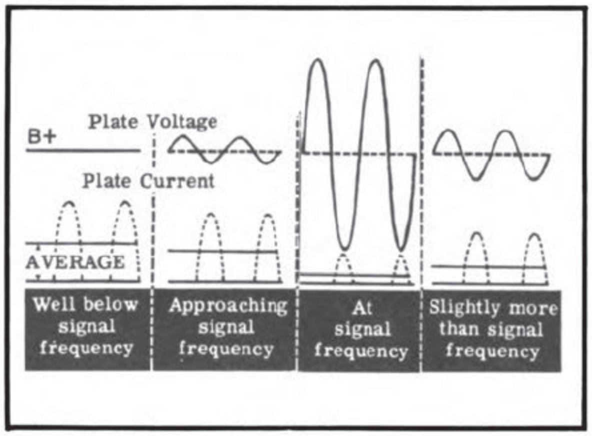


Figure 25-25 - Variations of plate voltage and current as tuning varies.

plate current meter, it is called tuning for a DIP.

In addition to the plate current meter, there is another meter which indicates correct tuning of the plate circuit. This meter is in the grid circuit of the following stage and is labeled M2 in Figure 25-26.

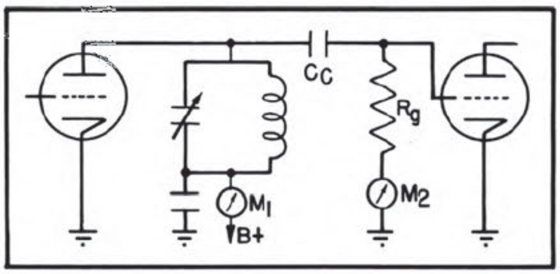


Figure 25-26 - Class C amplifier utilizing plate and grid current meters for tuning.

When the plate circuit is tuned to the frequency of the input signal, the voltage developed across the circuit is greatest and the output from that amplifier stage is greatest. The larger the output from that stage, the greater is the signal to the grid of the following stage.

The grid of the following stage will draw current whenever the input signal drives the grid positive. The larger the signal input, the greater will be the flow of current from the cathode to the grid. Since the signal input to the grid will be greatest when the plate circuit of the previous stage is accurately tuned, the grid will draw maximum current and milliammeter M2 (which measures the average grid current) will indicate a maximum reading. Thus when the plate tank is accurately tuned, the plate current meter indicates a DIP and the grid current meter of the following stage simultaneously registers a rise known as a PEAK reading.

If the grid circuit has fixed bias or combination bias, no grid current will be drawn until the signal is fairly large. This will happen some time after the plate current meter has started to dip. For this reason, the rising grid current indication is sharper than the decreasing plate current indication.

The normal procedure for tuning a stage which has a plate current meter and is followed by a stage which has a grid current meter, is to tune first for a minimum plate current. This indication is broader and less likely to be overlooked when varying the tuning. After observing the plate current starting to decrease, watch the grid current meter for a rise. The final adjustment will be for a rise in grid current. Since this is a sharper indication, tuning based on this indication will be more accurate.

Q11. Is maximum efficiency realized in a class C amplifier when the stage is tuned or detuned? Why?

FREQUENCY MULTIPLICATION

Up until now, it has been assumed that the plate tank circuit of a class C amplifier stage in a transmitter can be tuned only to the grid signal frequency, whatever that may be. For example, if the grid signal frequency is 1 Mc, the plate tank circuit is also tuned to 1 Mc. If the grid signal is a pure sine wave, (no harmonic content) the plate circuit can be tuned only to the frequency of this sine wave (called the fundamental) and none other. However, generated frequencies are very seldom pure; they usually contain harmonics of the fundamental frequency. For example, if the master oscillator (operating class C) generates a 1 Mc sine wave that sine wave is rich in harmonics. It contains not only the fundamental (1 Mc) but also the second harmonic (2 Mc), the third harmonic (3 Mc) etc. Therefore, if a signal rich in harmonics is placed on the grid of a tuned amplifier the plate tank circuit may be tuned to any one of the harmonics that is present in the original grid signal. The process by which the input frequency to the grid is converted to a higher one in the plate by tuning to a harmonic of the fundamental, is called FREQUENCY MULTIPLICATION. The need for frequency multiplier circuits in a transmitter result from the fact that the frequency stability of any oscillator is dependent upon its operating frequency. At the lower frequencies relatively good oscillator stability can be achieved, but as the frequency increases the interelectrode capacitance of the

A11. Tuned, because, at resonance the high impedance of plate tank causes the plate current to be at its minimum value. Low plate current means less heat dissipation, therefore, greater efficiency.

tube and the stray lead inductances cause the oscillator to drift from its assigned frequency. Therefore, to insure frequency stability in a transmitter, the master oscillator is operated at a low frequency and a series of frequency multipliers are incorporated to raise the transmitter frequency to the desired output frequency.

25-25. Operation

Because the grid driving signal of a class C amplifier has appreciable harmonic content it is possible to employ a RF amplifier as a frequency multiplier to double, triple or quadruple the input frequency. A basic multiplier circuit, operating as a frequency doubler is illustrated in Figure 25-27.

The frequency multiplier circuit is operated class C with the plate tank resonant to twice the grid signal frequency. A combination of fixed and grid leak bias is incorporated. Fixed bias for no signal tube protection and grid leak bias for amplitude stability. When the signal applied to the grid rises above the cutoff value of the tube, there will be a pulse of current at the same frequency as the input signal flowing from the cathode to the plate energizing the plate tank circuit. Since the pulse of plate current contains appreciable energy at the second harmonic and the resonant frequency of the plate tank circuit is determined by the values of L and C, the pulse of tube current excites the plate tank circuit and causes it to resonate at a frequency that is twice the grid signal frequency. When the tube goes into cutoff, the energy supplied to the plate tank circuit is sufficient to continue oscillations in the plate tank circuit between current pulses. The reason the tuned circuit continues to oscillate is that the pulses of current always arrive at the same time during alternate cycles of the doubled frequency, thus energizing the plate tank circuit at the right time. The circuit in Figure 25-27, may be operated as a frequency tripler by tuning the plate tank circuit to the third harmonic of the grid input signal. The pulses of plate current arrive at the tuned circuit during every third cycle of output voltage and deliver enough energy to the tuned circuit to sustain oscillations during those cycles when no current flows.

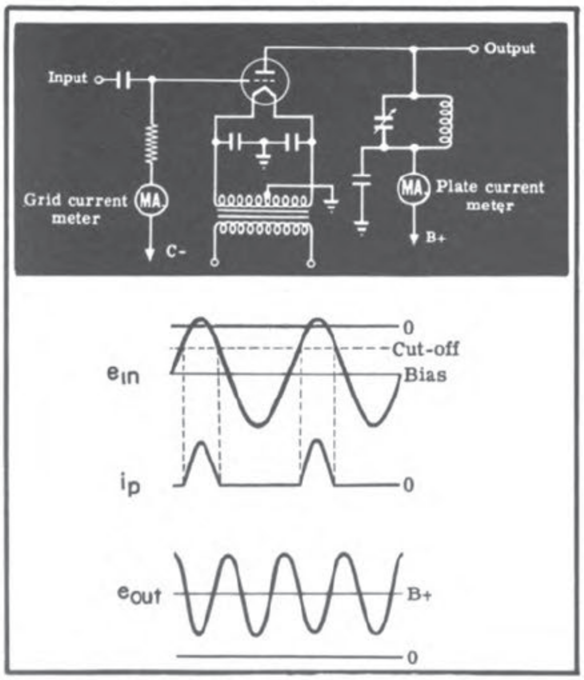


Figure 25-27 - Frequency doubler.

25-26. Conditions for Efficient Frequency Multiplication

The proper selection of the angle of plate current flow in a frequency multiplier, is a compromise between high efficiency and high power output. The curves of plate voltage, grid voltage and plate current of a class C amplifier are shown in Figure 25-28.

The dotted curves, indicate operation as a class C amplifier without frequency multiplication, and the solid curves, indicate operation as a frequency doubler. Operating a stage as a straight through amplifier, (no frequency multiplication) illustrates that plate current flows only when the plate voltage is at its minimum value, resulting in high efficiency. Operating the stage as a frequency doubler, with the same angle of plate current, results in plate voltage being higher at the beginning and end of the plate current pulse, perhaps even in excess

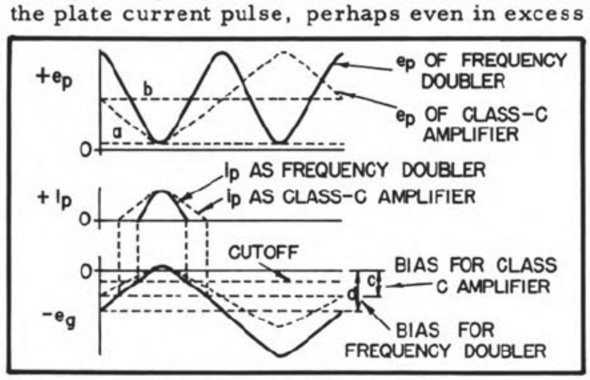


Figure 25-28 - Plate and grid voltages and plate current in a frequency doubler.

of the plate supply voltage, resulting in increased plate dissipation and a decrease in efficiency. To increase the efficiency of the stage when it is operating as a frequency doubler, the angle of plate current flow must be reduced. Decreasing the angle of plate current flow may be accomplished by increasing the bias. The shorter the length of current pulses, the higher plate circuit efficiency will be when generating a particular harmonic, however, the grid exciting voltage must be increased if appreciable power output at this harmonic is to be realized. The power output of a frequency multiplier varies inversely with the extent of frequency multiplication due to decreasing angle of plate current. Values for θ_p representing a practical compromise between high efficiency and high power output are given in Table 25-4.

Harmonic	θ_p	Percentage of Power Output
2	90^o - 120^o	65
3	80^o - 120^o	40
4	70^o - 90^o	30
5	60^o - 70^o	25

Table 25-4 - Angles of plate current and percentage of power output of frequency multiplier.

If the plate tank is tuned to the second harmonic of the grid tank, the duration of flow of plate current, ranges from 90^o to 120^o and the power output is about 65 percent of the output of a conventional straight through class C amplifier using the same tube. If the plate tank is tuned to the third harmonic of the grid tank, the angle of flow of plate current, ranges from 80^o to 120^o and the power output is reduced to 40 percent of that of a class C amplifier. If the plate tank is tuned to the fourth harmonic of the grid input signal, the angle of plate current flow is reduced to between 70^o and 90^o and the power output is 30 percent of that of a class C amplifier. If the frequency is multiplied by five, the angle of plate current flow ranges from 60^o to 70^o and the output power is 25 percent of that of a class C amplifier.

In every case it is necessary to increase the operating bias and the grid driving signal as the frequency multiplication increases in order not to overheat the triode plate. The flywheel effect in the plate tank supplies the missing cycles of grid drive and the output is approximately an undamped wave having sine waveform.

Three important conditions must prevail in order to obtain frequency multiplication—(1) high grid driving voltage, (2) high grid bias, and (3) plate tank circuit tuned to the desired harmonic. If the second harmonic is selected, the stage is called a FREQUENCY DOUBLER; if the third is used, the circuit is called a FREQUENCY TRIPLER.

25-27. Tuning Indications

All radio transmitters must be properly tuned to insure efficient operation on the assigned frequencies. Plate current meters are generally used to indicate proper adjustment of the RF stages. All stages, with the exception of the oscillator, are always adjusted or tuned for minimum plate current. If a stage is not tuned to resonance, the plate current will be high and high plate dissipation, power loss and low output will result.

To tune a stage to its assigned frequency (fundamental or harmonic), start by setting the tuning capacitor of the plate tank circuit for maximum capacitance. Maximum capacitance (plates meshed) means that the plate tank circuit is at its lowest frequency. Start decreasing the tank capacitance until a dip is noted in the plate current meter. The first dip indicates

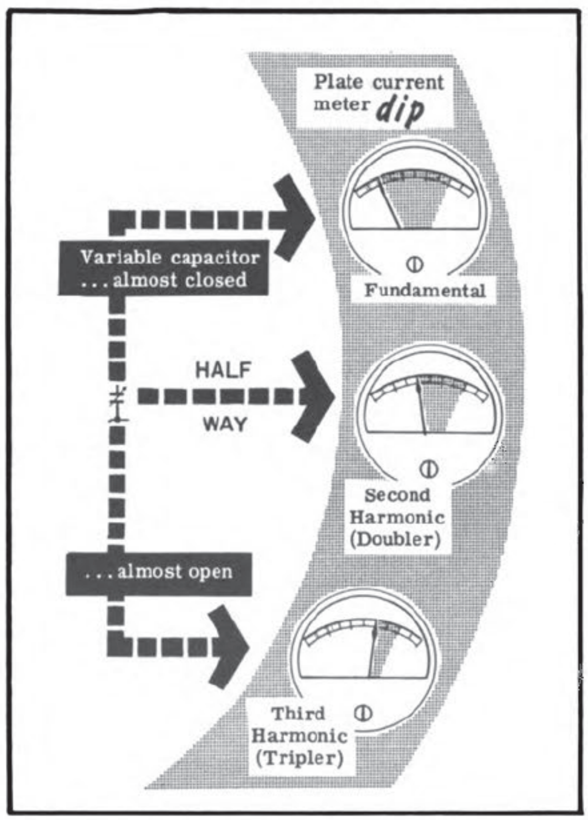


Figure 25-29 - Tuning to harmonics.

that the tank circuit is tuned to the fundamental. As you continue decreasing the capacity, you come to a second dip (not as pronounced as the first one) which is the second harmonic. Continue decreasing capacity and you may come to a third dip (provided the circuit constants are correct) which is not as pronounced as either the first or second dip. This dip indicates that the plate tuned circuit is tuned to the third harmonic.

Another method to tell to what frequency the plate tank circuit is tuned is to use a frequency indicator such as a wavemeter. A wavemeter is a device used to measure the frequency of a radio wave or electric oscillations. Absorption type wavemeters operate on the principle that the energy in an RLC circuit is absorbed by an adjoining RLC circuit. The amount of energy that is absorbed reaches a maximum value (coils fixed) when each circuit is resonant to the frequency in the tuned circuit under test, like the one shown in Figure 25-30.

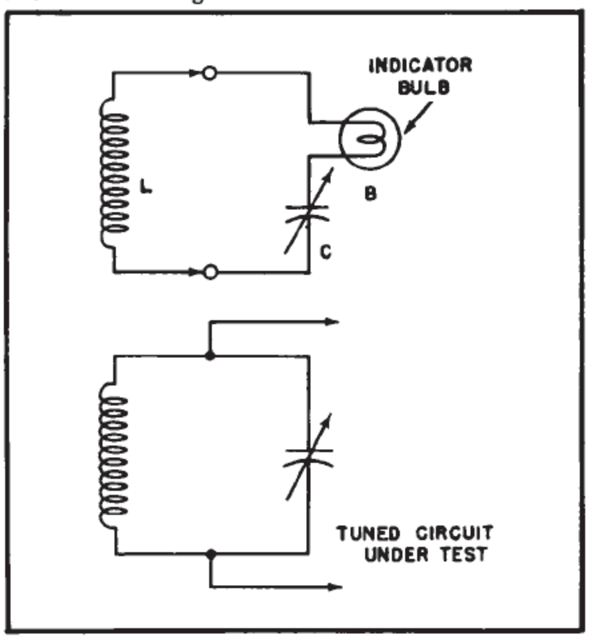


Figure 25-30 - Absorption wavemeter.

Maximum current is indicated in the wavemeter circuit by the indicator bulb glowing brightly. Therefore, the wavemeter is held in the vicinity of the circuit under test and the dial of the variable capacitor, C, is adjusted until the greatest brilliancy is noted on the indicator bulb. The value of resonant frequency is then determined from the dial on the variable capacitor.

Q12. An oscillator is operated at a frequency of two megacycles. If its output is applied to two stages of frequency tripling, what is the output frequency?

Q13. Could a push-pull amplifier be operated as a frequency doubler for more power output?

NEUTRALIZATION

In the fundamental RF amplifier, the input signal is applied to the grid tank circuit and the power output is taken from the tuned plate circuit. If a triode tube were utilized, the plate, cathode and other electrodes of the tube form an electrostatic system, each electrode acts as one plate of a small capacitor. Although the value of these interelectrode capacitances are not more than a few picofarads, they may have substantial effect on tube operation, especially at radio frequencies. For example, the low reactance of the grid to plate capacitance (C_{gp}), at radio frequencies, will provide an internal feedback path between the plate circuit and the grid circuit and if sufficient energy, having the same frequency and the same phase as the grid voltage, is fed back through this path, oscillations will occur. Although this type of internal feedback is frequently employed in oscillator circuits, it is undesirable in amplifier applications because it causes distortion, spurious radiations and interference to nearby radio receivers.

When triodes are used as RF amplifiers, it is possible to eliminate these oscillations by a process called NEUTRALIZATION. In neutralization, a network is included in the amplifier which feeds back to the grid, through an external circuit, a voltage which is at all times equal but opposite in phase to the voltage fed back to the grid through the plate to grid capacitance. Since the feedback voltage through C_{gp} is cancelled out by the voltage fed back through the external circuit, oscillations cannot take place. The use of a well shielded tetrode or pentode makes neutralization unnecessary, because the plate and grid are shielded from each other by the screen grid and its associated RF bypass capacitor which holds the screen at RF ground potential. However, the overall efficiency of these tubes is not as great as that of triodes, since there is a screen grid power loss.

There are several well known neutralization systems in use. Two of these, the PLATE or HAZELTINE neutralization system, and the GRID or RICE system, have the advantage of being useful over a wide frequency range and derive their names from the part of the circuit in which the feedback voltage is developed.

25-28. Plate Neutralization (Hazeltine)

The method of neutralization, most frequently used, plate neutralization, is shown in Figure 25-31. This is a typical transformer coupled RF amplifier to which has been added a neutralizing inductance L_2 closely coupled to L_1 and a neutralizing capacitor C_N. L_2 is connected in such a manner that the polarity of voltage at point B of this transformer is in phase opposition to the voltage at the corresponding end point A. The center of this transformer (point C) is placed at RF ground through the low reactance of the RF bypass capacitor C_4. Therefore, the RF voltages measured at points A and B with respect to ground are 180 degrees out of phase and equal in amplitude (assuming point C is the exact center). C_{gp} is the grid-to-plate internal capacitance, represented in the schematic as a capacitor external to the tube. C_N is the neutralizing capacitor, that is, the capacitor through which the neutralizing signal is coupled to cancel the effects of C_{gp}.

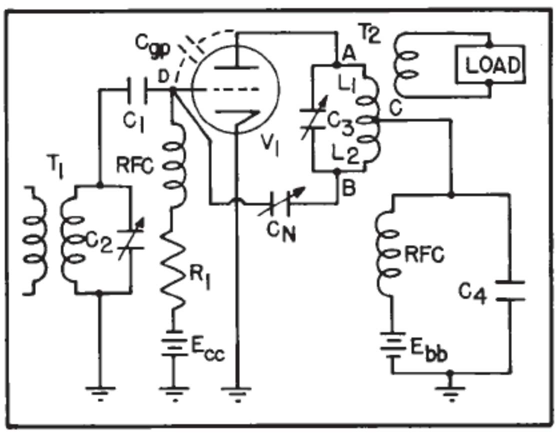


Figure 25-31 - Plate neutralization circuit.

The operation of the plate neutralization circuit can best be understood with the aid of the equivalent circuit of Figure 25-32. Point C of the plate tank circuit is effectively the same point as the bottom of the grid tank through the low reactance of C_4. The low reactance of the coupling capacitor C_1 makes the top of the grid tank effectively the same as point D. Therefore, the plate neutralization circuit resolves down to a bridge arrangement consisting of the grid tank circuit, L_1 and L_2, and C_{gp} and C_N.

When the potential across the plate tank appears as in Figure 25-32 (positive at the top and negative at the bottom) currents will be caused to flow in the directions indicated. When the plate tank voltage reverses its polarity, the direction of current flow will be opposite to the direction indicated. Assuming the potentials generated by L_1 and L_2 are equal and the values

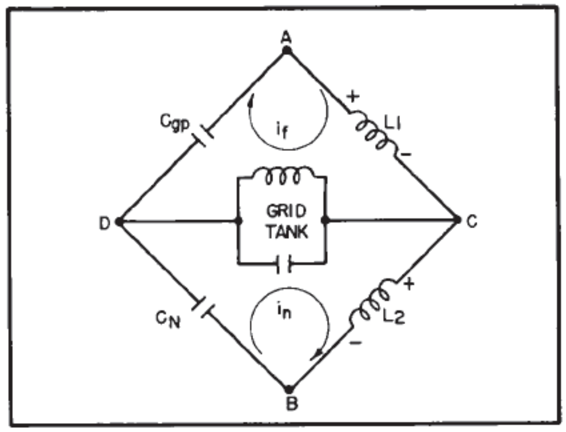


Figure 25-32 - Equivalent circuit.

of capacitances C_{gp} and C_N are equal, the two currents will also be equal. When these currents are equal, there will be no difference in potential between points C and D and, therefore, no resultant current will flow through the grid tank circuit. No current flow through the input tank (in the equivalent circuit) indicates that no energy has been fed back from the output to the input, thus, no regeneration or degeneration can take place and the circuit is considered to be neutralized. If the neutralizing capacitor C_N were smaller in value than C_{gp}, its higher reactance would cause i_f to be the predominate current resulting in a feedback current through the input tank circuit which is regenerative. This regenerative signal would result in circuit oscillations. If the neutralizing capacitor were larger in value than C_{gp}, its lower reactance would cause i_n to be the predominate current. This current through the grid input tank would develop a feedback voltage that is degenerative. This degenerative signal will result in reduced output from the stage. Therefore, proper neutralization of a RF amplifier stage is realized when the feedback voltage through the neutralizing capacitor (C_N) cancels the feedback voltage through C_{gp} resulting in no energy transfer from the output circuit to the input circuit.

25-29. Grid Neutralization (Rice)

Another circuit which provides a means of neutralizing the effects of the grid-to-plate capacity is the grid neutralization circuit shown in Figure 25-33. It differs from plate neutralization in that the split tank circuit, which provides the neutralizing voltage, is located in the grid circuit. The operation of this circuit can best be demonstrated with the aid of an equivalent circuit as shown in Figure 25-34. Triode tube can be represented by an equivalent gen-

A12. Eighteen megacycles.

A13. No. Push-pull amplifier cancels even harmonics.

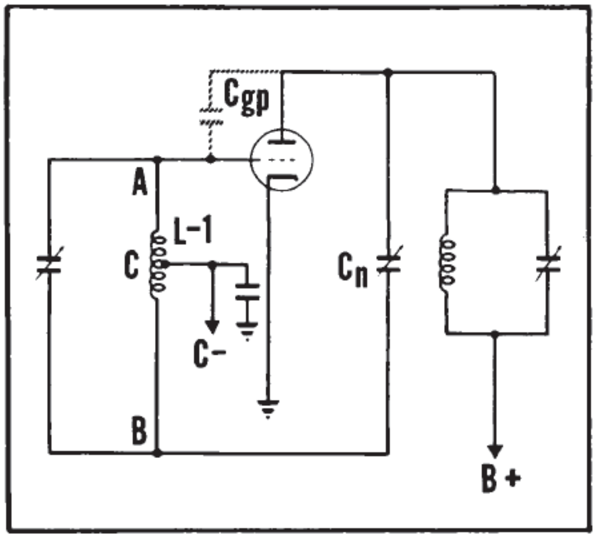


Figure 25-33 - Grid neutralization.

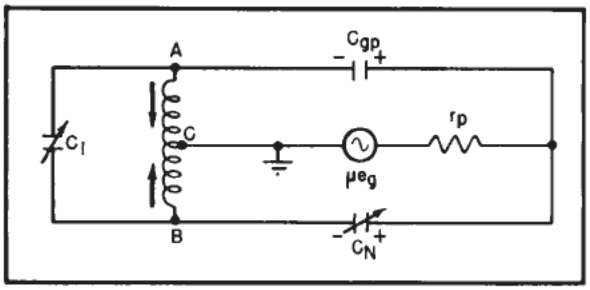


Figure 25-34 - Equivalent circuit.

erator (μe_g) and its plate resistance (r_p). C_{gp}, the interelectrode capacitance of the tube, is located between the plate and grid which is effectively point A, the top of the grid tank. C_N, the neutralizing capacitor, is connected between the tube plate and point B, the bottom end of the grid tank. The low reactance of C_2 places point C of the grid tank at RF ground. Initially, before the application of an input signal, C_{gp} and C_N will charge with the polarities indicated. Their charge path is from C- through the respective capacitors to the B+ supply. With the application of an input signal, when the top of the grid tank goes positive (point A), tube current increases and the increased voltage drop across the plate tank causes the plate voltage to decrease in value. As plate voltage decreases, Cgp and C_N will commence to discharge through the grid tank. If $C_N = C_{gp}$, their discharge currents through L_1 will set up opposing fields across L_1 and effectively neutralize each other. If the feedback current through C_{gp} were greater, the resultant field would develop a feedback voltage across the grid tank which would be regenerative resulting in circuit oscillations. Therefore, the prevention of oscillations can be achieved by insuring that the feedback through C_N counteracts the feedback through C_{gp}.

25-30. Coil Neutralization

Coil neutralization differs from the systems previously described. In the other circuits, the voltage fed back through the plate to grid capacitance is canceled by an equal but opposite voltage fed back through an external path (C_N). In this circuit (Figure 25-35), the effect of plate to grid capacitance can be nullified by paralleling it with an inductor having the same value of reactance. Since the two reactances are equal and opposite, a parallel resonant circuit exists between the grid and the plate. Since a parallel resonant circuit exists, there is no transfer of energy through the circuit from plate to grid. C_1 is a blocking capacitor which prevents the plate supply voltage from being felt on the control grid of the tube.

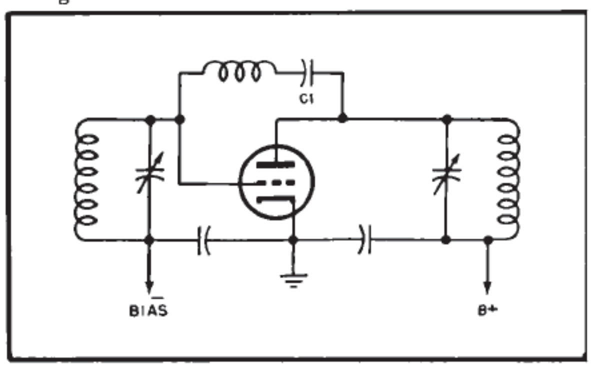


Figure 25-35 - Coil neutralization.

25-31. Neutralizing Procedure

The procedures for neutralizing are almost independent of the type of neutralizing circuit used. At the start of neutralization, the plate voltage is removed from the stage to be neutralized so that any signal present in the plate circuit is due to the interelectrode capacity coupling between the grid and plate. Then the master oscillator and those amplifier stages which precede the unneutralized stage are tuned. This will provide a strong signal to the grid of the unneutralized stage. The next step depends on the indicator used but it always results in the adjustment of the neutralizing capacitor (C_N) until there is a minimum amount of energy transferred to the plate circuit.

Place a pick up coil near the plate tank and connect it to the vertical input of a cathode ray oscilloscope. C_N is then adjusted so that no RF voltage appears on the scope when the plate tank is tuned to resonance. Under these circumstances, the RF current divides equally through C_{gp} and C_N. The resulting RF currents in the plate tank flow in opposite directions and cancel the tank inductive effect so that no resonant build-up occurs between the coil and capacitors. A neon glow lamp, a loop of wire attached to the filament connections of a flashlight bulb, or a sensitive RF galvanometer may be used if an oscilloscope is not available.

If there is a milliammeter in the amplifier grid circuit, the adjustment of C_N may be made by observing the grid meter as the plate tank is tuned through resonance, with no plate voltage applied. When there is an unbalance between C_{gp} and C_N, the plate becomes alternately positive and negative as the plate tank approaches resonance. On positive swings, plate current flows. As the plate tank circuit is tuned to the resonant frequency, some of the electrons that were going to the grid now go to the plate, thereby causing a dip in the grid current. However, if C_N is adjusted to neutralize the amplifier stage, the RF current from the input stage divides equally and flows in opposite directions in the two halves of the plate tank coil, thus canceling the inductive effect of the coil and preventing the build-up of resonance in the tank. There is no rise in tank current and voltage, and the triode plate remains at zero potential. Therefore, with C_N properly adjusted, no dip in grid current occurs as the plate tank is tuned through the resonant frequency.

In some transmitter circuits, it is more convenient to turn off the filament voltage on the amplifier stage instead of removing plate voltage. If this is done, the process of neutralizing the amplifier is carried out in the same way, except that no current flows in the amplifier grid circuit. The absence of radio frequencies in the amplifier plate tank, is evidence of the correct adjustment of C_N.

Once a neutralizing capacitor is adjusted for a particular tube, it will require only occasional checks. However, if the tube is changed for a new one, the neutralizing capacitor will need adjustment since the new tube may have a slightly different value of C_{gp}.

Q14. Is it necessary to neutralize a stage which is operated as a frequency doubler?

25-32. Parasitic Oscillations

Parasitic oscillations are oscillations at some frequency usually far removed from the frequency to which the transmitter is tuned. Any inductor will resonate at some frequency when associated with a capacitance. Occasionally, various transmitter components which possess both inductive and capacitive properties will cause the circuit to oscillate at their common resonant frequency. The inductance may be that of wiring, leads of capacitors, a section of a coil or RF choke, or the element leads within a tube. The capacitance may be that of normal circuit capacitors, or the capacitance between turns of a coil or choke, or the interelectrode capacitance of the tube. Parasitics usually are eliminated in the design of the transmitter but they sometimes appear after the set has been modified or if some parts are replaced. Defective tubes are another cause of parasitics. The presence of parasitic oscillations is indicated by a rough, nonmusical note in the receiver and an indication of plate and grid current in a properly neutralized amplifier when excitation is removed. Parasitics reduce the useful power output of the transmitter by absorbing some of the power which should be useful output. They may cause excessive current that blow fuses, trip overload relays, ruin capacitors and inductors in the oscillating circuit, and damage the tubes.

High frequency parasitics usually can be removed by inserting small RF chokes or resistors in series with each grid and plate connection. These should be placed as close as possible to the tube terminals. Chokes for parasitic suppression have very low inductance and negligible distributed capacitance. The resistor can be approximately 50 ohms. An efficient parasitic suppressor can be made by winding a coil of wire on the body of a small carbon resistor and connecting the coil and resistor in parallel. This combination is usually most effective in grid circuits but its use may be necessary in some plate circuits. The presence of the parasitic suppressor in grid circuits makes the amplifier harder to drive at high frequencies but the decrease in the power sensitivity is compensated for by the lack of spurious oscillations. Low-frequency parasitics occur most often in amplifiers having RF chokes in both grid and plate circuits. Sometimes the tube or tuning capacitor may be tapped down on a tank coil to provide proper impedance matching and to insure maximum energy transfer at the desired frequency.

USE OF RF POWER AMPLIFIERS

25-33. Buffers

One of the characteristics of class C amplifiers is the consumption of power in the input circuit caused by the grid voltage going positive

A14. No. Since the frequency in the plate circuit of a frequency doubler differs from the grid signal frequency, the feedback voltage through C_{gp} to the grid will not result in circuit oscillations.

with relation to the cathode and drawing current during part of the operating cycle. This power, called EXCITATION POWER, must be supplied by the preceding stage in the transmitter. For the power amplifier to operate efficiently, a minimum amount of excitation is required. This is determined by the type of tube and the dc voltage applied to it. The stages prior to the power amplifier must be able to supply this excitation without overloading. Because the frequency of an oscillator depends to some extent on the load impressed upon it, it is undesirable from the standpoint of frequency stabilization to attempt to supply excitation power directly from the oscillator. Another consideration is the modulation impressed on a carrier wave in the power amplifier. In this case, a varying amount of excitation is demanded by the power amplifier as the modulation changes. This changing load also can seriously affect the frequency stabilization of the oscillator if the oscillator is used to drive such a modulated amplifier directly.

Therefore, a BUFFER amplifier is introduced between the oscillator and the power amplifier to isolate the two stages from each other. The buffer amplifier is usually operated class A, so that it will not affect the oscillator. In this condition, no power is drawn in its grid circuit. For class A service, the efficiency is low, and tubes of fairly high ratings must be used in buffer circuits for high-power final amplifiers. In broadcasting service, many buffer stages are used to build up the low-level output from the oscillator to a value sufficient to provide excitation to the power amplifier. In general, the buffer must supply from 5 to 20 percent as much power as the final amplifier will produce.

The buffer amplifier must supply this excitation and have considerable reserve power so that its output does not vary with changing load. This is termed GOOD REGULATION. Since the efficiency of the class A buffer is low, its plate dissipation can be as much as one-half that of the tube used as the final power amplifier. Excitation requirements increase as the frequency of the operation is increased. This is because losses in the input circuit are greater at higher frequencies.

25-34. Dummy Antenna

A dummy antenna is a device which has the necessary power handling capabilities and impedance characteristics of an antenna but does not radiate or receive radio signals. A resistor in series with an ammeter connected across the output terminals of a transmitter provides a load for the final power amplifier and all the power from the final power amplifier is dissipated in this load.

TROUBLESHOOTING TRANSMITTER

25-35. The Oscillator

The purpose of the electron-coupled master oscillator is to generate a stable RF signal which can be varied over a given range.

The ECO operates as follows: The oscillator section of the ECO is composed of the grid and screen circuits and is a Colpitts oscillator. The oscillator frequency is determined by the grid-screen tank circuit consisting of L_1, C_1, C_2 and C_3. The screen, which acts as the plate of the oscillator section, is coupled to the tank circuit through the RF bypass capacitor, C_5. Grid-leak bias is developed across R_1 by the discharge of C_4. The RF choke in the cathode circuit provides a low resistance DC path to ground for the cathode. However, the high reactance of the choke to RF does not allow RF to flow through it. The RF must flow through C_3 (the feedback capacitor) to the cathode. The screen dropping resistor, R_2, drops the screen voltage to the correct value. The RF oscillations generated in the oscillator section of the ECO are electron-coupled to the plate through the flow of plate current. The RF choke in the plate lead acts as a high impedance for the RF signal and serves the same purpose as the plate load resistor in an audio amplifier. The RF coupling capacitor, C_6, passes the signal to the signal to the grid of the IPA.

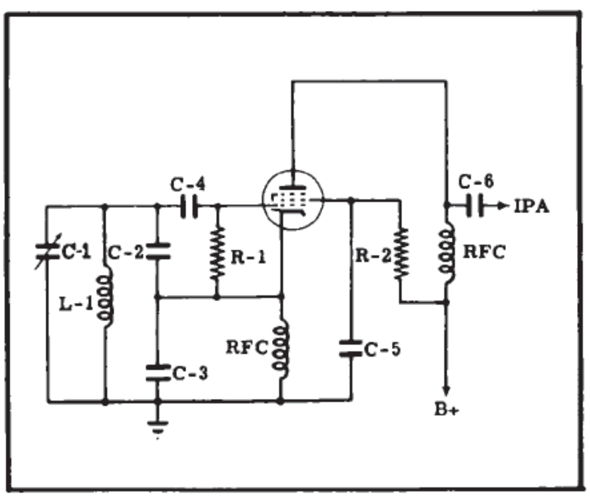


Figure 25-36 - Master oscillator (ECO).

25-36. The Intermediate Power Amplifier

The purpose of the intermediate power amplifier is to isolate the oscillator for improved frequency stability and to amplify the RF signal in order to drive the power amplifier efficiently. The IPA also serves to increase the tuning range, if desired, by doubling or tripling the generated frequency in its plate tank circuit.

The operation of the IPA is essentially as follows: A combination of grid-leak and cathode bias is provided by R_3, C_6 and R_4, C_7 respectively. Resistor R_5 drops the screen voltage to the correct value. The screen by-pass capacitor, C_8, is returned directly to the cathode rather than to ground. This provides a more direct path back to the cathode for any RF variations on the screen. The RF coil in the plate lead acts as a high impedance for the RF signal and serves the same purpose as the plate load resistor in an audio amplifier. C_9 is a coupling capacitor which passes the RF to the tank circuit and at the same time blocks the DC. The plate tank circuit, C_{10} and L_2, can be tuned to the IPA grid signal, in which case the IPA is said to operate "straight through," or the tank circuit can be tuned to twice the grid signal frequency, and in this case the IPA is called a "doubler." When the IPA doubles, the isolation between the grid and plate circuits is improved and as a result there is less chance of the IPA breaking into oscillation. Doubling has another advantage in that it raises the carrier frequency while permitting the oscillator to operate at a lower frequency where it will be more stable. Capacitor C_{11} couples the RF to the grid of the power amplifier.

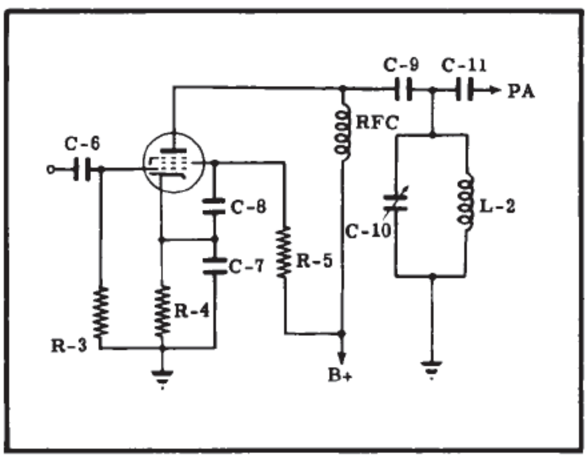


Figure 25-37 - Intermediate power amplifier (IPA).

25-37. The Power Amplifier

The purpose of the power amplifier is to increase the power of the RF signal so that it can be radiated by the antenna. The PA usually operates straight through for good efficiency. Only in unusual cases does the PA act as a doubler.

The PA operates as follows: Capacitor C_{11} couples the RF from the output of the IPA to the grid of the PA. Here as in the IPA there is a combination of grid-leak and cathode bias provided by R_6 and C_{11}; and R_7 and C_{12}, respectively. The RF choke while providing a DC path from plate to B+ also acts as a high impedance plate load for the RF signal. C_{13} couples the RF to the tuned circuit and blocks the DC.

The plate tank circuit C_{15}, L_3 is tuned to the grid signal frequency and a high RF voltage is developed across it. The high powered RF signal in the plate tank is coupled by coil L_5 to the antenna for radiation. Coil L_4 couples some energy back to the grid through capacitor C_{14}, called a "neutralizing capacitor."

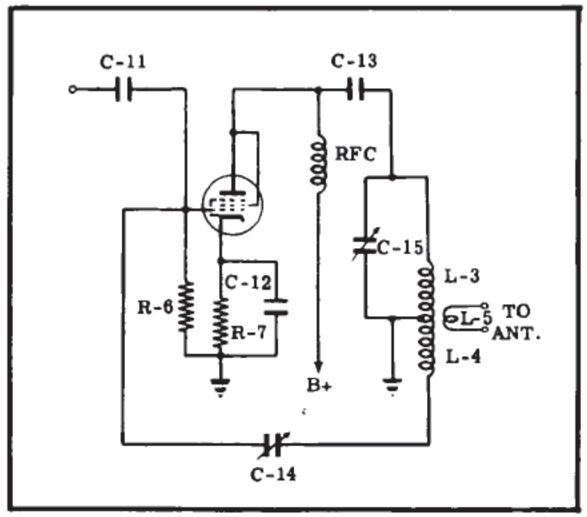


Figure 25-38 - Final power amplifier (FPA).

25-38. The Troubleshooting Method for Transmitters

The troubleshooting method involving visual checks and signal tracing, which you have employed to troubleshoot power supplies, amplifiers and oscillators, is not used as such when troubleshooting Navy transmitters. Because of the high voltages present in transmitters, it is not safe to work with them with the power on. Therefore an alternate troubleshooting method is provided by means of meters mounted on the transmitter's front panel. These front panel meters are permanently wired into various transmitter circuits, and most troubles can be located using only these meters. The only other meter required is an ohmmeter which is used to make resistance checks (with power off) after the trouble has been localized to a particular stage by means of front panel meter troubleshooting.

In the table below are listed normal indications for typical front panel meters. Since any determination of a malfunction will first be indicated by front panel meter indications a thorough understanding of normal readings is necessary. The normal indications listed apply to successive circuits and to the portion for which they are listed.

Meter	Normal Indications	Immediate Conclusions
DC Grid Current	Current rises to normal value when plate circuit of previous state is tuned	1. RF signal present on the grid 2. Coupling circuit OK 3. DC grid circuit complete to ground 4. Cathode emitting
DC Plate Current	Current dips to normal value when plate circuit is tuned	1. RF present in plate circuit 2. DC plate circuit complete to power supply 3. Tube operating normally 4. Cathode emitting

25-39. Troubleshooting Using Front Panel Meter Readings

Below is a troubleshooting chart for a transmitter which lists abnormal front panel meter readings and their possible causes. Any one of the items in the column of possible troubles can cause the abnormal meter reading.

Meter	Symptoms	Possible Trouble
DC Grid Current	Zero reading	1. No grid drive 2. No RF in plate of previous stage 3. Bad tube
	Low reading	1. Weak grid drive 2. Mistuned plate circuit of previous stage 3. Bad tube
	High reading	1. Open screen dropping resistor (no screen voltage) 2. No plate voltage
DC Plate Current	Zero reading	1. Bad tube 2. No grid drive (fixed bias, below cut-off) 3. Open screen dropping resistor (fixed bias, below cut-off) 4. No plate voltage 5. Open filament return to ground
	Low reading	1. Weak tube 2. Open grid leak resistor
	High reading	1. No grid drive (no fixed bias) 2. Mistuned plate circuit

The way to use the chart is to be aware of all the possible troubles that could cause an abnormal meter reading. Then by observing other meter readings, you eliminate possible troubles until you are left with one or two. Now you can de-energize the transmitter and troubleshoot the suspected circuits with an ohmmeter until you have uncovered the defect.

25-40. Safety Precautions

Transmitter tubes in Navy equipment operate at very high plate voltages and it is, therefore, dangerous to reach inside a transmitter to measure currents or to change a part while the power is on. This may seem obvious to you but unless you are always aware of the danger, an accident is possible.

Navy transmitters are designed to protect you from accidents that could happen through carelessness. They have access doors through which you can remove bad tubes. If you were to open one of these doors when the transmitter was on, the power would turn off automatically. The doors operate switches, called INTERLOCKS, which are connected between the on-off switch and the circuits of the transmitter. When

an access door is open, the interlock is open and no power gets to the transmitter. Thus, opening the access door shuts down the transmitter. Because of the important role played by an interlock, shorting out an interlock is dangerous and forbidden.

There is still possibility of a serious shock—even when you do turn the power off before reaching into a transmitter—a fault in the circuit may have prevented a capacitor from discharging and your hand will provide a discharge path.

As a safeguard against this, a shorting bar, consisting of a metal rod with a wooden handle, is used. The metal end is connected with a copper braid to the ground of the transmitter. Before putting your hand inside, you probe about with the shorting rod, touching every point which might conceivably be charged. If nothing is charged, this procedure costs you about 30 seconds; if something is charged, the shorting rod discharges it and the procedure saves you from a severe jolt.

KEYING METHODS

Keying a transmitter causes an RF signal to be radiated ONLY when the key contacts are closed. When the key is open the transmitter does not radiate energy. Keying is accomplished in either the oscillator or amplifier stages of a transmitter. A number of different keying systems are used in Navy transmitters.

In most Navy transmitters the hand telegraph key is at low potential with respect to ground. The keying bar is usually grounded to protect the operator. Generally a keying relay with its contacts in the center tap lead of the filament transformer is used to key the equipment. Because one or more stages use the same filament transformer, these stages are also keyed. The class C final amplifier, when operated with fixed bias is usually not keyed because with no excitation applied no current flows. Hence, keying the final amplifier along with the other stages is not necessary.

25-41. Oscillator Keying

Two methods of oscillator keying are shown in Figure 25-39. In Figure 25-39A, the grid circuit is closed at all times, and the key opens and closes the negative side of the plate circuit. This system is called PLATE KEYING. When the key is open, no plate current can flow and the circuit does not oscillate. In Figure 25-39B, the cathode circuit is open when the key is open and neither grid current nor plate current can flow. Both circuits are closed when the key is closed. This system is called CATHODE KEYING.

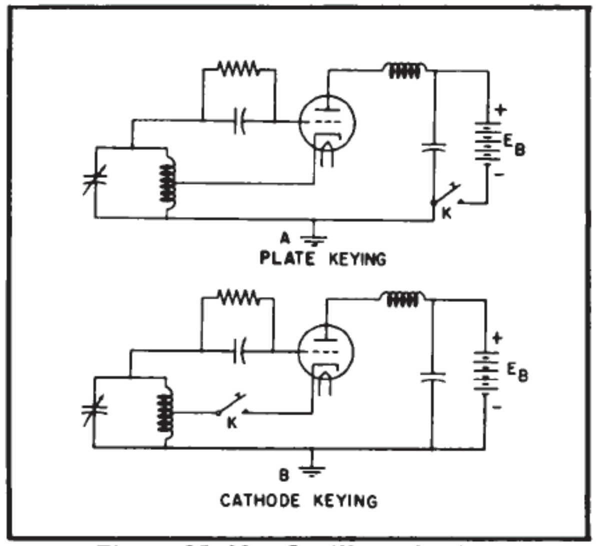


Figure 25-39 - Oscillator keying.

Although the circuits of Figure 25-39 may be used to key amplifiers, other keying methods are generally employed because of the larger values of plate current and voltage encountered.

25-42. Blocked Grid Keying

Two methods of blocked grid keying are shown in Figure 25-40. The key in Figure 25-40A, shorts cathode resistor R_1, allowing normal plate current to flow. With the key open, reduced plate current flows up through resistor R_1, making the end connected to grid resistor

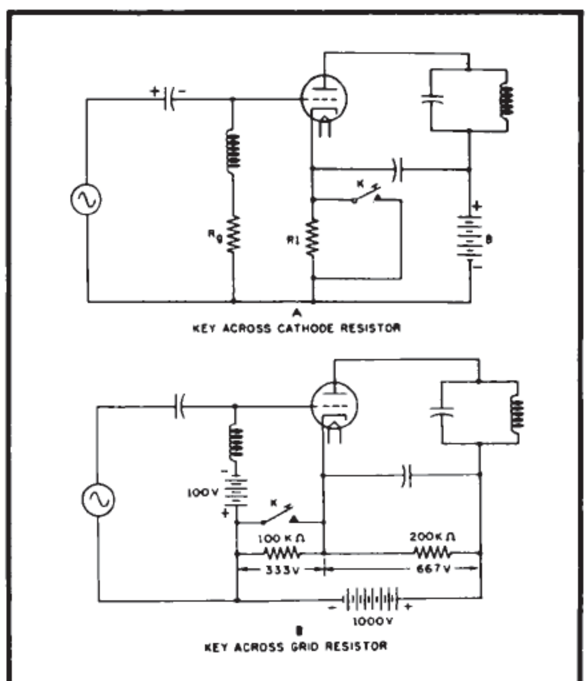


Figure 25-40 - Blocked grid keying.

R_g negative. If R_1 has a high enough value, the bias developed is sufficient to cause cutoff of plate current. Depressing the key short circuits R1, thus increasing the bias above cutoff and allowing the normal flow of plate current. Grid resistor R_g is the usual grid leak resistor for normal bias. This method of keying is applied to the buffer stage in a transmitter.

The blocked grid keying method shown in Figure 25-40B, affords complete cutoff of plate current and is one of the best methods for keying amplifier stages in transmitters. In the voltage divider, with the key open, two thirds of 1,000 volts, or 667 volts, are developed across the 200 k-ohm resistor and one third of 1,000 volts, or 333 volts, are developed across the 100 k-ohm resistor. The grid bias is the sum of -100v and -333v, or -433 volts. Because this is below cutoff, no plate current flows. The plate voltage is 667 volts. With the key closed the 100 k-ohm resistor is shorted out and the voltage across the 200 k-ohm resistor is increased to 1,000 volts. Thus, the plate voltage becomes 1,000 volts at the same time the grid bias becomes -100 volts. Grid bias is now above cutoff and the amplifier triode conducts. Normal amplifier action follows.

Where greater frequency stability is required, the oscillator should remain in operation continuously while the transmitter is in use. This procedure keeps the oscillator tube at normal operating temperature and offers less chance for frequency variation to occur each time the key is closed. If the oscillator is to operate continuously and the keying is to be accomplished in an amplifier stage following the oscillator, the oscillator circuit must be carefully shielded to prevent radiation and interference to the operator while he is receiving.

25-43. Keying Relays

In transmitters using a crystal controlled oscillator, the keying is almost always in a stage following the oscillator. In the large transmitters (75 watts or higher) the ordinary hand key cannot accommodate the plate current without excessive arcing. Moreover, because of the high plate potentials used it is dangerous to operate a hand key in the plate circuit. A slight slip of the hand below the key knob might result in a bad shock; or, in the case of defective RF plate chokes, a severe RF burn might be incurred. In these larger transmitters, some local low voltage supply, such as a battery or the filament supply to the transmitter, is used with the hand key to open and close a circuit through the coils of a keying relay. The relay contacts in turn open and close the keying circuits of the amplifier tubes. A schematic diagram of a typical relay-operated keying system is shown in Figure 25-41. The hand key closes the circuit from the low voltage supply through coil (L) of the keying relay. The relay armature closes the relay contacts as a result of the magnetic pull exerted on the armature. The armature moves against the tension of a spring. When the hand key is opened, the relay coil is deenergized and the spring opens the relay contacts.

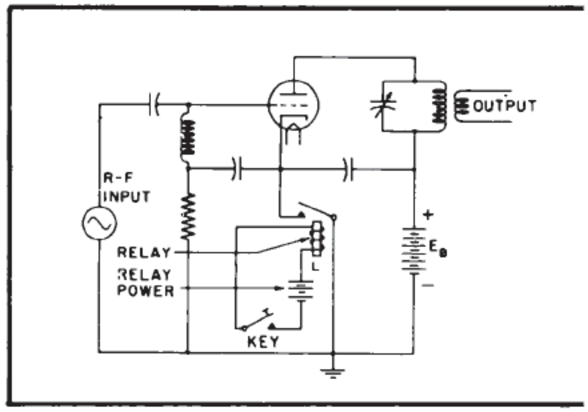


Figure 25-41 - Circuit for a relay operated keying system.

25-44. Key Clicks

Theoretically, keying a transmitter should instantly start and stop radiation of the carrier completely. However the sudden application and removal of power creates rapid surges of current which cause interference in nearby receivers. Even though such receivers are tuned to frequencies far removed from that of the transmitter, interference is present in the form of clicks or thumps. To prevent such interference, key click filters are used in the keying systems of radio transmitters. Two types of key click filters are shown in Figure 25-42.

The capacitors and RF chokes in both circuits of Figure 25-42 prevent surges of current. The choke coil, L, causes a lag in the curren

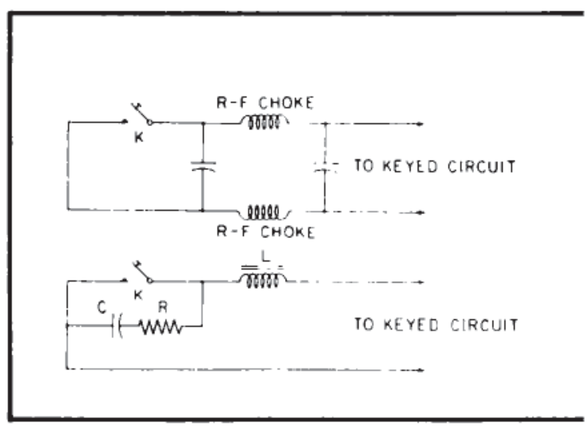


Figure 25-42 - Key-click filters.

when the key is closed, and the current builds up gradually instead of instantly. Capacitor C charges up as the key is opened and slowly releases the energy stored in the inductor magnetic field. Resistor R controls the rate of charge and discharge of capacitor C and also prevents sparking at the key contacts by the sudden discharge of C when the key is closed.

Another difficulty that may be encountered in keying a transmitter is the presence of a back wave. A back wave results when some RF energy leaks through to the antenna even though the key is open. The effect is as though the dots and dashes were simply louder portions of a continuous carrier. It may be difficult to distinguish the dots and dashes under such conditions. Back wave radiation is usually the result of incomplete neutralization.

EXERCISE 25

1. What is the difference between an audio power amplifier and an RF power amplifier?
2. What class of operation is usually used for an RF power amplifier?
3. Compare the power output and plate efficiency of amplifiers operated class B and class C.
4. What type of bias is usually used in an RF power amplifier? Why?
5. Describe the relationship between the current waveforms in an RF power amplifier.
6. What causes the dip in the plate current waveform?
7. What sort of voltage waveform may be expected from a class C RF power amplifier with an inductive load?
8. What sort of output voltage waveform may be expected if a tank circuit is the plate load?
9. What is the phase relationship between plate voltage and plate current in an RF power amplifier?
10. What is the relationship between the time duration of a plate waveform of voltage, and a plate waveform of current?
11. What is meant by the term "grid current loading?"
12. What is exciting power?
13. How are average constant current curves plotted?
14. What is the difference between an operating line and a load line?
15. What is the relationship between the angle of grid current and the angle of plate current?
16. What is a harmonic?
17. Describe the type of tubes usually used for RF power amplifiers.
18. Why are directly heated filaments preferred?
19. Describe the various methods of coupling, their advantages and disadvantages.
20. What is the difference between a shunt-fed and a series-fed amplifier?
21. Compare the response curves for a single-tuned and a double-tuned RF amplifier.
22. Describe the amounts of coupling and how the shape of the response curve is affected by changing the Q.
23. On what principle does the frequency multiplier operate?
24. What is neutralization?
25. Why is neutralization necessary?
26. What is the difference between grid and plate neutralization?
27. Primarily, what must the technician look for when tuning a CW transmitter?
28. What is the main objection to the use of cathode keying?
29. What is meant by the term "key clicks."
30. Why are relays sometimes used in keying circuits?

CHAPTER 26

AMPLITUDE MODULATION

Long range radio communications sprang into existance in December 1901 when Guglielmo Marconi, an Italian physicist, succeeded in receiving, at St. Johns Newfoundland, radio signals transmitted across the Atlantic Ocean from Cornwall, England. From this first transmission, radio communications have evolved into the sophisticated systems presently used to communicate with space vehicles millions of miles from the earth.

To be of practical value the radio signal must carry some type of intelligence. This intelligence is applied to the radio signal at the sending end of the system, and is removed and utilized at the receiving end. In a radio communications system, the signal which carries the intelligence is a high frequency sine wave called a CARRIER wave. The intelligence which is applied to the carrier wave is called MODULATION and can be in the form of a code, voice signals, music, or pictures. When intelligence has been applied to a carrier, the carrier is said to be MODULATED.

One of the first methods used to modulate a carrier wave consisted of turning the carrier on and off in accordance with a predetermined code. Each letter of the alphabet is represented by a sequence of dots and dashes. To send a dot, the carrier is turned on for a brief instant. To send a dash, the carrier is turned on for a longer period of time. Thus, one by one, each letter of the message is transmitted to the receiver. A radio signal of this type is called a CONTINUOUS WAVE (CW) signal. Although CW signals are one of the oldest forms of modulation, they are still used extensively in Navy communications.

In many instances communications could be carried on more efficiently if voice signals could be transmitted rather than code. This would extend the use of radio communications to those persons not trained to send and receive code. The following section describes several methods used to voice modulate a carrier wave.

26-1. Types of Modulation

When the modulating signal is combined with the carrier wave in such a manner as to produce NEW frequencies on either side of the carrier frequency, the process is called modulation. Although a number of systems have been developed to produce modulation, only three of these systems are commonly used for radio communications. In one such system, the new frequencies which appear as a result of modulation, when added to the carrier frequency, produce a resultant waveform which rises and falls in amplitude. Because of the amplitude variations which occur in the resultant wave as a result of modulation, this system of modulation is called AMPLITUDE MODULATION (abbreviated AM).

In another system of modulation, the modulating signal combines with the carrier in such a way as to cause the frequency of the resultant wave to vary in accordance with changes in the instantaneous amplitude of the modulating signal. This system of modulation is called FREQUENCY MODULATION (FM).

In the third system, the modulating signal controls the phase of the resultant wave thereby causing PHASE MODULATION (PM).

This chapter will describe the methods used to apply voice signals to a carrier wave by the process of amplitude modulation (AM).

26-2. Typical AM Transmitter

An AM transmitter can be divided into two major sections according to the frequencies at which they operate. One section is called the RF unit and is the section of the transmitter used to generate the radio frequency carrier wave. As illustrated in Figure 26-1, the carrier originates in the oscillator stage where it is generated as a constant amplitude—constant frequency sine

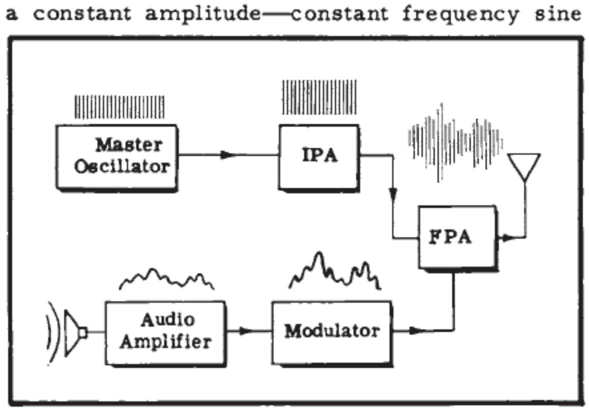


Figure 26-1 - Block diagram of AM transmitter.

wave. As it leaves the oscillator stage, the carrier is not of sufficient amplitude and must pass through one or more stages of amplification before it attains the high power required by the antenna. With the exception of the last stage, the amplifiers between the oscillator and the antenna are called INTERMEDIATE POWER AMPLIFIERS (abbreviated IPA). The last or final stage which connects to the antenna is called the FINAL POWER AMPLIFIER (FPA).

The second section of the transmitter contains the audio circuitry. This section of the transmitter takes the minute signal from the microphone and increases its amplitude to the amount necessary to fully modulate the carrier. The last audio stage applies its signal to the carrier and is called the MODULATOR. (See Figure 26-1. Thus, intelligence is included in the radiated wave.

26-3. Heterodyne Process

In previous chapters great emphasis was placed on linear operation of any circuit in which an audio signal was present. Operating potentials were carefully chosen so that the signals utilized only the straight-line or linear portion of the tube's characteristic curve. When operated in this manner, the output signal voltage is identical to the input signal voltage in both waveshape and frequency content.

If the operating point of the tube is shifted to the lower knee of the dynamic transfer curve, the output signal from the amplifier will be distorted. If the signal applied to the amplifier contains more than one frequency, the individual frequencies will interact with each other and new frequency components will be generated. Thus, when an amplifier is operated in a non-linear manner, THE OUTPUT SIGNAL CONTAINS FREQUENCIES WHICH ARE NOT PRESENT IN THE INPUT SIGNAL.

Of the various new frequencies generated when two single frequencies are introduced into a non-linear amplifier, two of these frequencies are more prominent than the others. One of these important new frequencies has a frequency equal to the SUM of the two original frequencies, while the second new frequency is equal to the DIFFERENCE between the two original frequencies. The process of combining two frequencies in such a way as to produce sum and difference frequencies is called MIXING or HETERODYNE ACTION.

In Figure 26-2, the results of both linear (Q_1) and non-linear (Q_2) operation of a vacuum tube are illustrated. As shown in the inset, the signal applied to the grid of the tube is the sum of a 5kc sine wave (E_1) and a 100kc sine wave (E_2). When this complex signal is applied to the grid of the tube at operating point Q_1, all

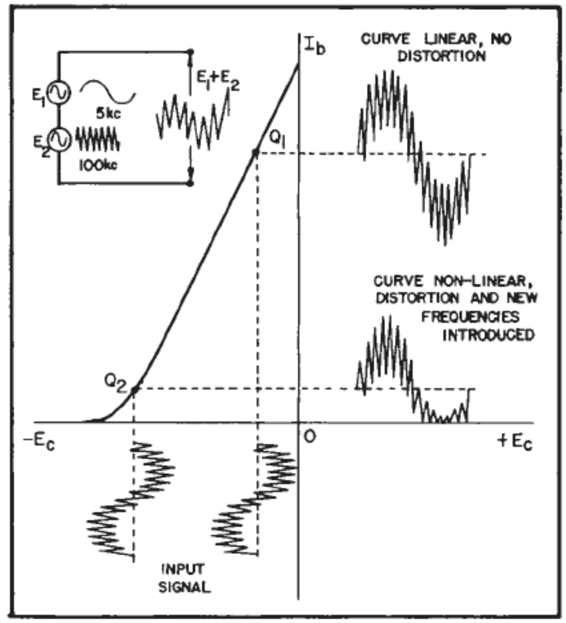


Figure 26-2- Heterodyne action occurs as a result of non-linear operation.

parts of the input signal operate on the linear section of the transfer curve. Since over this section of the curve plate current is proportional to grid voltage, the output waveform is an enlarged replica of the input waveform. This being true, the output waveform must contain the same two frequencies that were applied to the input, and no other frequencies.

If the tube is operated at point Q_2 and the same two-frequency input signal is applied to the grid, the output waveform will be severly distorted as shown in Figure 26-2. Notice, that the less negative portions of the input signal receive greater amplification than do the more negative portions of the input signal.

The presence of distortion in the output signal indicates that new frequencies are generated by the tube and added to the two original frequencies applied to the grid. Indeed, if filter circuits are connected to the tube, the various frequency components can be separated and viewed individually.

A complete analysis of the output waveform for operating point Q_2 shows that the output signal contains harmonics of each of the two original frequencies, plus, the sum and difference frequencies mentioned previously. Because amplitude modulation is concerned mainly with the sum and difference frequencies, the harmonic frequencies will not be considered further.

In Figure 26-2 the input frequencies are 100kc and 5kc. Due to heterodyne action the

output signal will contain 105kc (100kc + 5kc) and 95kc (100kc - 5kc), in addition to the two original frequencies of 100kc and 5kc. It should be pointed out that all of these frequencies, are radio frequencies capable of being radiated into space from a transmitter antenna.

When the human voice, or sounds from a musical instrument are converted into electrical impulses by a microphone, the electrical impulses will contain a large number of frequencies. If this audio signal is then mixed with an RF signal (such as 100kc) a large number of sum and difference frequencies will be generated ONE SET OF SUM AND DIFFERENCE FREQUENCIES WILL BE PRODUCED FOR EACH AUDIO FREQUENCY THAT HETERODYNES WITH THE RF SIGNAL.

Q1. Why does an audio amplifier, such as a public address amplifier, not introduce sum and difference frequencies into the audio output?

26-4. The Modulated Wave

The frequencies present in a signal can be conveniently represented by a graph of the FREQUENCY SPECTRUM. In this graph, shown in Figure 26-3, each individual frequency is portrayed as a vertical line. The position of the line along the horizontal axis indicates the frequency of the signal, while the height of the line is proportional to the amplitude of the signal. The RF spectrum in Figure 26-3 shows the frequencies present when heterodyning occurs between frequencies of 100kc and 5kc.

For reasons to be discussed in a later chapter, it is not practical to radiate energy at audio frequencies. The heterodyne principle makes possible the conversion of an audio signal into an RF signal which can be radiated or transmitted through space.

Referring again to Figure 26-3, the sum and difference frequencies are located very near the RF signal, while the audio signal is spaced a considerable distance away. Because of this frequency separation, the audio frequency can be easily removed by filter circuits, leaving three radio frequencies of 95kc, 100kc, and 105kc. It is these three radio frequencies which are radiated through space to the receiving station. At the receiver the process is reversed. The frequency of 95kc, for example, is heterodyned with the frequency of 100kc and sum and difference frequencies are again produced. (A similar process occurs between the frequencies of 100kc and 105kc). Of the resultant frequencies (95kc, 100kc, 195kc, and 5kc) all are filtered out in the receiver with the exception of the 5kc difference frequency. This frequency, which is identical to the original 5kc audio signal applied to the transmitter, is retained and amplified. Thus, the 5kc audio tone APPEARS to have been radiated through space from the transmitter to the receiver.

In the process just described the 100kc frequency is called the carrier frequency and the sum and difference frequencies are called "side frequencies." Since the sum frequency appears above the carrier frequency it is called the UPPER SIDE FREQUENCY. The difference frequency appears below the carrier and is called the LOWER SIDE FREQUENCY.

When a carrier is modulated by voice or music signals, a large number of sum and difference frequencies are produced. All of the sum frequencies above the carrier are spoken of collectively as the UPPER SIDE BAND, while all the difference frequencies below the carrier, taken as a group, are called the LOWER SIDE BAND.

If the carrier and the modulating signal are constant in amplitude the sum and difference

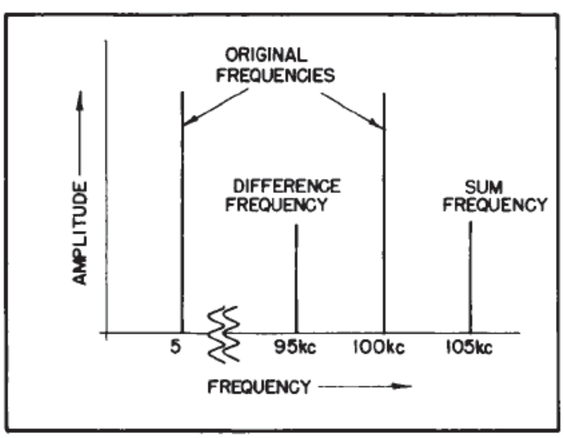


Figure 26-3 - Radio frequency spectrum.

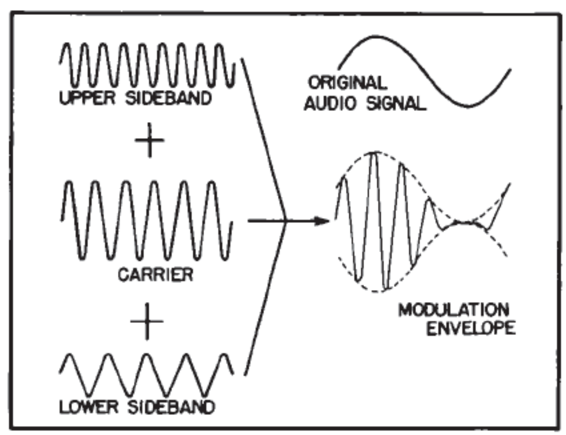


Figure 26-4-Formation of the modulation envelope.

A1. All stages in an amplifier of this type are operated on the linear portion of their characteristic curves, thus, heterodyning does not occur to any great extent.

---

frequencies will also be constant in amplitude. However, when the carrier and sidebands are combined in a single impedance and viewed simultaneously with an oscilloscope, the resultant waveform appears as shown in Figure 26-4. This resultant wave is called the MODULATION ENVELOPE. The modulation envelope has the same frequency as the carrier, but rises and falls in amplitude as the continual phase shift between the carrier and sidebands causes these signals to first aid, and then oppose one another. These cyclic variations in the amplitude of the envelope have the same frequency as the audio modulating voltage. It was this characteristic variation in amplitude which gave rise to the descriptive, but oversimplified and erroneous concept of the audio intelligence "riding" or being "superimposed" on the carrier. The audio intelligence is actually contained in the spacing or difference between the carrier and sideband frequencies.

26-5. Bandwidth of AM Wave

An ideal carrier wave contains but a single frequency and occupies a very minute section of the frequency spectrum. When the carrier is modulated, sideband frequencies are created above and below the carrier frequency, causing the signal to use up a greater portion of the frequency spectrum. The amount of space in the spectrum required by the signal is called the BANDWIDTH of the signal.

The bandwidth of a modulated wave is a function of the frequencies contained in the modulating signal. For example, when a 100kc carrier is modulated by a 5kc audio tone, sideband frequencies are created at 95kc and 105kc. This signal requires 10kc of space in the spectrum.

If the same 100kc carrier is modulated by a 10kc audio tone, sideband frequencies will appear at 90kc and 110kc and the signal will have a bandwidth of 20kc. Notice, that as the modulating signal becomes higher in frequency the bandwidth required becomes greater. From the two examples cited above, it can be seen that at any instant, the bandwidth of an amplitude modulated wave is two times the highest modulating frequency applied at that time. Thus if a 400kc carrier is modulated with 3kc, 5kc, and 8kc simultaneously, sideband frequencies will appear at 392kc, 395kc, 397kc, 403kc, 405kc, and 408kc. This signal extends from 392kc to 408kc and has a bandwidth of 16kc, twice the highest modulating frequency of 8kc.

Musical instruments produce complex sound waves, containing a great number of frequencies. The frequencies produced by a piano, for example, range from about 27 cycles per second to about 4200 cycles per second, with harmonic frequencies extending beyond 10kc. To transmit a musical passage with a high degree of fidelity, modulating frequencies up to 15kc must be included in the signal. This requires a bandwidth of at least 30kc to prevent attenuation of higher order harmonic frequencies.

If the signal to be transmitted contains voice frequencies only, and fidelity is of minor importance, the bandwidth requirements are far less stringent. A baritone voice encompasses frequencies of about 100 cycles per second to about 350 cycles per second. Intelligible voice communications can be carried out as long as the communications system retains audio frequencies up to several thousand cycles per second. Comparing the conditions for transmitting voice signals with those for transmitting music shows that much less spectrum space is required for voice communications.

Radio stations in the standard broadcast band are assigned carrier frequencies by the Federal Communications Commission (FCC). To prevent the sideband frequencies of one station from interfering with those of an adjacent station, the carriers of the two stations must be spaced a sufficient distance apart in the radio spectrum. The standard broadcast band starts at 535kc and ends at 1605kc. Carrier assignments start at 540kc, and continue in a succession of 10kc increments until the upper limits of the broadcast band are reached. This adds up to a total of 107 carrier assignments or CHANNELS over the entire broadcast band. If stations were assigned to all 107 channels (in a given geographical area) each station would be alloted a channel width of 10kc. This leaves 5kc on each side of each carrier for sidebands. Since it would be nearly impossible to prevent stations so closely spaced from interfering with one another, the FCC avoids assigning adjacent ahannels to stations in the same area. As a consequence of this policy, one or more vacant channels normally exist between stations in the broadcast band. In the interest of better fidelity the stations are permitted to use modulating frequencies higher than 5kc (even though sideband frequencies are produced in the vacent adjacent channels), as long as no interference with other stations is produced.

Q2. Why would it be desirable to suppress the higher audio frequencies in a communications transmitter?

26-6. Basic Modulator Circuit

Modulation of one signal by another can be accomplished in ANY non-linear device. Thus, modulation could be produced in any type of vacuum tube such as a diode, triode, etc., a transistor, or even in a transformer, if the transformer is driven into partial saturation. For reasons of efficiency and minimum distortion, transmitter modulator circuits normally utilize only vacuum tube or transistor amplifiers as the element in which the heterodyning occurs.

In the vacuum tube modulator the audio intelligence and the radio frequency carrier can be applied to the same element of the tube, or to separate elements of the tube. In most modulator circuits the RF carrier is applied to the control grid, and the audio modulating voltage is then applied to the control grid or other tube element such as the plate, screen grid, cathode, etc. The name given to the method of modulation depends on the element to which the modulating signal is applied. If the modulating signal is applied to the plate of the tube, the method is called PLATE MODULATION. Thus, the common methods of modulation are: PLATE MODULATION, CONTROL GRID MODULATION, SCREEN GRID MODULATION, SUPPRESSOR GRID MODULATION, and CATHODE MODULATION. A modulation method also exists in which the modulating signal is applied to both screen grid and plate.

The schematic diagram of a plate modulation system is illustrated in Figure 26-5. The neutralization circuit and decoupling networks have

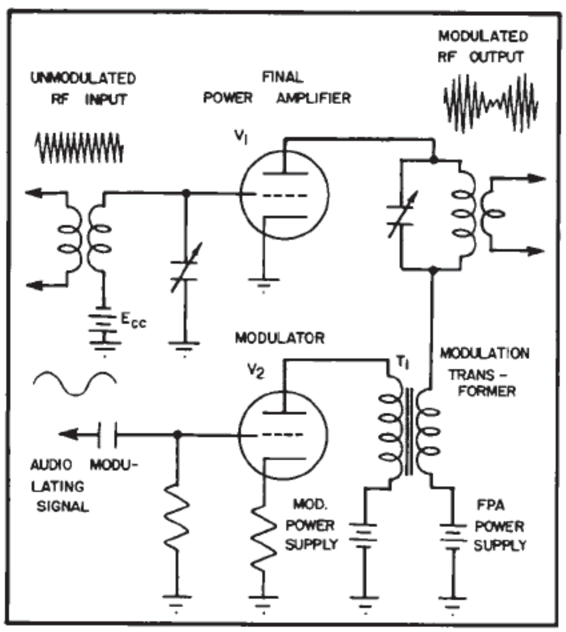


Figure 26-5 - Circuit for producing plate modulation.

been omitted from the diagram for simplicity. Stage V_1 is a final RF power amplifier and is normally operated class C to introduce non-linearity and to take advantage of the high efficiency of this class of operation. The unmodulated carrier from the intermediate power amplifier section of the transmitter is applied to the grid of V_1, and the amplitude modulated RF signal is inductively coupled from the plate tank coil of the stage.

The modulating signal (the intelligence) is introduced into the plate of V_1 by modulation transformer T_1. The secondary of this transformer is connected in series with the plate supply voltage to the final so that the audio signal across the secondary of T_1 forms part of the final amplifier plate supply voltage. Since part of the final plate voltage is supplied by the modulator, a portion of the plate input power to V_1 is obtained from the modulating signal. Thus, the modulator is a power amplifier stage and the final RF power amplifier acts as the load on the modulator. Transformer T_1 provides impedance matching between the modulator and the final RF amplifier. In some low power transmitters the modulator may obtain its plate voltage from the same power supply as the final, rather than from a separate supply as shown.

PLATE MODULATOR ANALYSIS

In order to show how plate modulation is brought about, the operation of a typical, medium power, plate modulation circuit will be analyzed. The tube selected for this example is an 812-A, for which the tube manual lists the following characteristics when used as a class C plate-modulated RF power amplifier:

Typical Operation:

DC Plate Voltage	1000	volts
DC Grid Voltage	-110	volts
Peak RF Grid Voltage	220	volts
DC Plate Current	115	ma
DC Grid Current (approx.)	33	ma
Driving Power (approx.)	6.6	watts
Power Output (approx.)	85	watts

It will be assumed that the modulated stage in Figure 26-5 has the characteristics listed above under "Typical Operation."

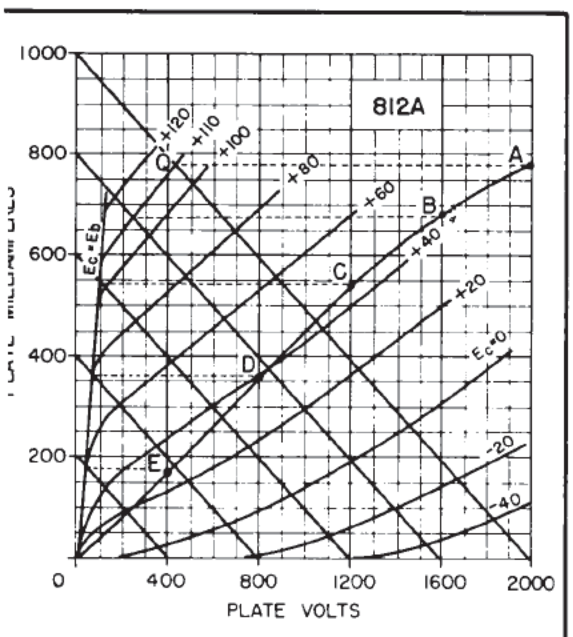
26-7. Transfer Curve

When the modulating signal is applied to the plate of the final RF power amplifier, the plate-to-cathode voltage of this amplifier will swing above and below E_{bb} in accordance with the modulating voltage. One method that can be used to investigate the modulated output signal consists of constructing a curve which shows

2. This reduces the bandwidth of the signal, permitting a greater number of stations to operate in a given frequency band.

ıe peak plate current as a function of plate upply voltage. The peak amplitude variations f the RF plate current pulses can then be plotted as the modulating signal swings the supply oltage above and below its unmodulated value. ince this curve is obtained by transferring oints from a load line, it can be considered as type of transfer characteristic. However, due › its limitations it is more correctly termed "quasi-transfer" curve.

Because the load impedance and total grid oltage are also paramount factors in determinıg the magnitude of plate current, these two uantities must be treated as constants if the urve is to represent the effects of plate voltge alone. In view of this, the plate tank ciruit will be assumed to present a resistive imedance of 2000 ohms to the final RF amplifier. o simulate a constant grid voltage, the curve ill be constructed for the instant of time when ıe RF carrier wave applied to the grid reaches s most positive value. According to the tube ıanual data the recommended bias for an 812-A s -110 volts and the peak RF grid voltage should e 220 volts. Thus, at the peak of each posive alternation of the carrier the grid will be riven to +110 volts (-110 + 220 = +110). Since ıe curve is constructed using this peak grid



igure 26-6 - Construction of the transfer curve.

voltage, the values of plate current obtained from the curve represent peak values of current.

In order to plot the quasi-transfer curve, plate current values must be obtained for a selected set of plate supply voltages. This is accomplished by constructing a load line for each of the supply voltages chosen, as shown in Figure 26-6. For example, at a plate supply potential of 2000 volts the load line extends from the 2000 volts point on the horizontal axis, to the 1000 milliamperes point on the vertical axis. At the positive peak of any given cycle of the grid signal, the grid voltage is +110 volts. Therefore, at this instant the operating point (Q) is located at the intersection of the load line and the +110 volt curve. This operating point shows that for a plate supply voltage of 2000 volts the peak plate current will be approximately 780 ma. These values of supply voltage and plate current are plotted as point A in Figure 26-6.

By a similar process load lines are constructed for plate supply potentials of 1600 volts, 1200 volts, 800 volts, and 400 volts. From these load lines, points B, C, D, and E are established. The transfer characteristic is then completed by drawing a smooth curve through these five points. Notice, that except for the uppermost part, the curve is very linear indicating that peak plate current varies directly with supply voltage.

The flattening of the quasi-transfer curve above point C occurs because the tube becomes unsaturated at this value of plate supply voltage. Below point C the grid signal is large enough to drive the operating point to the line representing $E_b = E_c$, causing complete saturation. At approximately 1200 volts supply voltage the grid signal is no longer large enough to produce saturation and the peak amplitude of the plate current pulses no longer has the same direct relationship to plate supply voltage as exists below point C.

26-8. Circuit Operation

The operation of a plate modulation circuit similar to the one shown in Figure 26-5 will be explained using the transfer curve developed in Figure 26-6. For convenience, both the circuit and the curve have been included in Figure 26-7. In accordance with the values recommended by the tube manual, the bias for V_1, the stage to be modulated, is set at -110 volts and the dc plate supply potential is 1000 volts. If a 100kc carrier having a peak amplitude of 220 volts is applied to the grid of V_1, each positive alternation of the carrier will bring the tube out of cut-off and produce a high amplitude pulse of plate current. When no modulating signal is applied to V_2 the plate supply of V_1 remains

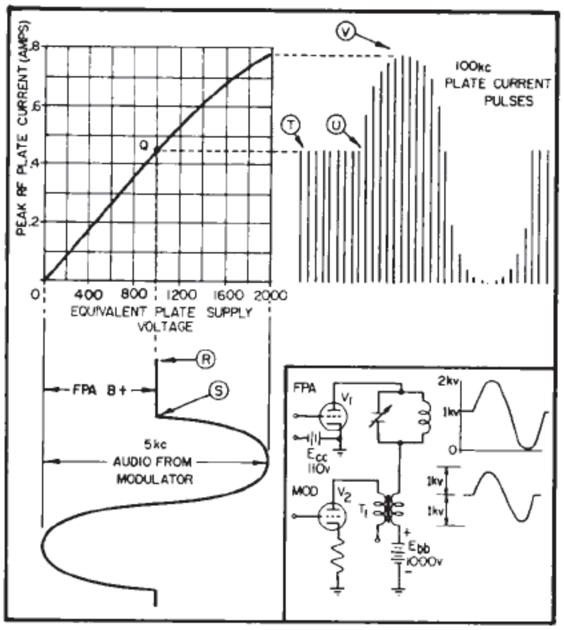


Figure 26-7 - Determination of plate current pulses from the transfer curve.

constant at 1000 volts (point R to point S) and the operating point of the tube is fixed at point Q. As long as the supply voltage remains at this value, a series of plate current pulses having a peak amplitude of approximately 450 ma will be generated (point T to point U).

At point S an audio sine wave of 5000 cps is applied to the grid of modulator tube V_2. This signal is amplified and appears across the secondary of modulation transformer T_1. As the positive alternation of the modulating signal begins at point S, the equivalent plate supply voltage starts to increase. During the time that the modulating voltage increases the equivalent supply voltage from 1000 to 2000 volts, the amplitude of the plate current increases as shown between point U and point V. Thus, for each new value of equivalent plate supply voltage produced by the modulating signal, a corresponding value of plate current can be obtained from the transfer curve. Notice, that the peak amplitude of the plate current pulses follows the modulating signal changes very closely. Thus, when the modulating signal has a peak amplitude equal to the FPA B+ the equivalent plate supply voltage will vary between 0 volts and $2E_{bb}$ and the plate current pulses will vary between zero and nearly two times their unmodulated value.

26-9. Plate Circuit Waveforms

After the amplitudes of the plate current pulses have been determined, the tank and plate voltage waveforms can be constructed. These relationships are shown graphically in Figure 26-8. Part A of this illustration shows the results of adding the audio modulating signal to the constant dc output voltage from the FPA power supply. The total voltage thus varies between zero and 2000 volts in accordance with the modulating signal.

When this voltage is applied to the plate of the final power amplifier, the plate current pulses are varied in amplitude as shown in part B of Figure 26-8. During the time the modulating signal increases the final power amplifier plate voltage to 2000 volts, the plate current pulses are very large in amplitude. At the negative peaks of the modulating signal the plate voltage on the final power amplifier is reduced to zero and no plate current pulses occur at this time.

As each pulse of plate current passes through the tank circuit connected in series with the

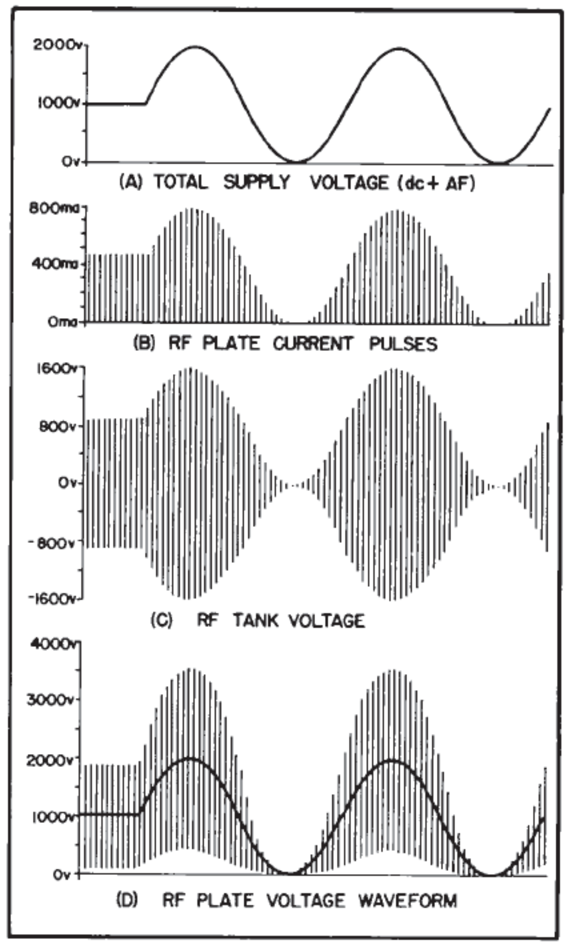


Figure 26-8 - Current and voltage waveforms in a class C plate-modulated RF power amplifier.

plate of the final power amplifier, a quantity of energy is applied to the tank. When the amplitude of the current pulse is large, a substantial amount of energy is added to the tank causing it to oscillate strongly. At times when the plate current pulse is small or entirely absent, little or no energy is supplied to the tank and the oscillations become weak and die out. In this manner a modulation envelope like the one shown in part C of Figure 26-8 is developed across the plate tank circuit.

The voltage waveform which exists across the final power amplifier tube is shown in part D of Figure 26-8. Each time the carrier voltage applied to the grid brings the tube out of cut-off, a narrow high amplitude pulse of plate current is produced, provided sufficient plate voltage exists across the tube. This pulse of plate current must flow through the impedance of the plate tank circuit, thereby producing a voltage drop across the tank which subtracts from the supply voltage. Since the grid signal is large enough to drive the grid highly positive, each plate current pulse causes nearly all of the supply voltage to drop across the tank impedance. As shown in the diagram, the negative peak of each RF voltage cycle lowers the plate voltage to about ten percent of the supply voltage. For example, before the application of the modulating signal in part D of Figure 26-8 the supply voltage is a constant 1000 volts and the plate voltage dips to approximately 100 volts on each negative alternation of the RF plate signal.

On each positive alternation of the plate signal, the RF sine wave across the tank adds to the supply voltage and the maximum plate voltage becomes nearly twice the value of the supply voltage. From this it follows that during the positive peak of the modulating signal, the total supply voltage (dc + AF) doubles causing the maximum RF plate voltage to approach four times the value of the dc plate supply voltage. In Figure 26-8 this amounts to approximately 3600 volts with a dc supply voltage of 1000 volts.

ANALYSIS OF AN AM WAVE

A significant amount of information concerning the basic principles of amplitude modulation can be obtained from a study of the properties of the modulation envelope. As pointed out previously, a carrier wave which has been modulated by voice or music signals is accompanied by two sidebands, teeming with individual frequencies which fluctuate continuously. Since a wave of this nature is nearly impossible to analyze, it will be assumed in the following sections that the modulating signal, unless otherwise qualified, is a single frequency, constant amplitude sine wave.

26-10. Percent of Modulation

The depth or degree of modulation is defined in terms of the maximum permissible amount of modulation. Thus, a fully modulated wave is said to be 100 PERCENT MODULATED. The modulation envelope in Figure 26-9 shows the conditions for 100% sine wave modulation. For this degree of modulation the peak audio voltage must be equal to the dc supply voltage to the FPA. Under this condition the RF output voltage will reach zero on the negative peak of the modulating signal and will rise to two times the amplitude of the unmodulated carrier on the positive peak of the modulating signal.

When analyzed, the modulation envelope in part A of Figure 26-9 is found to consist of a carrier and two sideband frequencies as shown in part B of this Figure. Since for 100% modulation the peak audio modulating voltage is approximately equal to the peak RF voltage, the combined sideband voltage is equal to the carrier voltage. Because the sideband voltage is divided between two sideband frequencies, AT 100% MODULATION EACH SIDE FREQUENCY HAS AN AMPLITUDE EQUAL TO ONE-HALF THE AMPLITUDE OF THE CARRIER.

If the audio modulating voltage is increased beyond the amount required to produce 100% modulation, the negative peak of the modulating signal becomes larger in amplitude than the

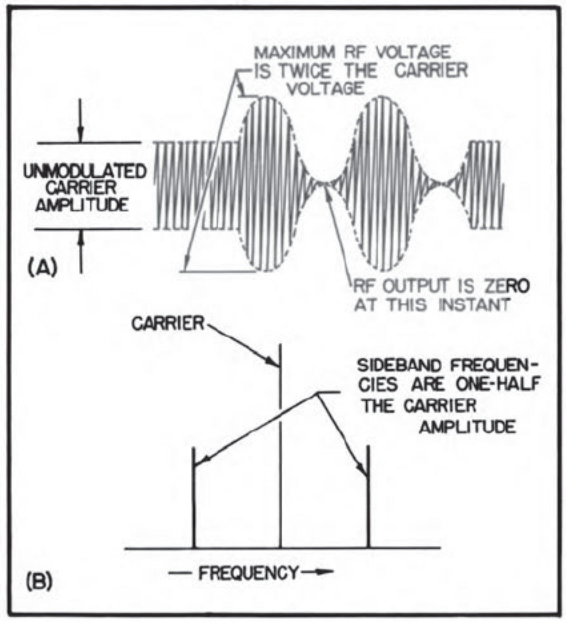


Figure 26-9 - Conditions for 100% modulation.

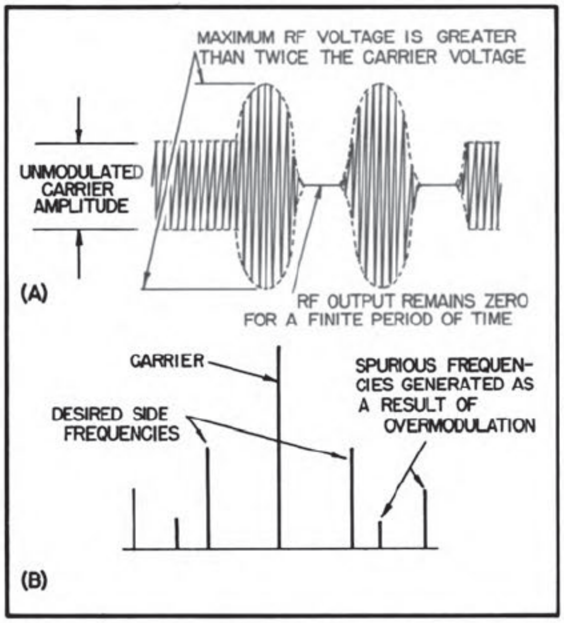


Figure 26-10 - Interference and distortion produced as a result of overmodulation.

dc plate supply voltage to the final power amplifier. This causes the final plate voltage to be negative for a short period of time near the negative peak of the modulating signal. For the duration of this negative plate voltage no RF energy is developed across the plate tank circuit and the RF output voltage remains at zero for this period of time, as shown in part A of Figure 26-10.

A careful examination of the modulation envelope in Figure 26-10A shows that the negative peak of the modulating signal has effectively been clipped. When detected (demodulated) in the receiver, the signal would have an appearance somewhat similar to a square wave. To the ear the signal would sound severely distorted (although this would depend on the degree of overmodulation).

If a radio receiver is tuned to a frequency near, but somewhat outside the channel on which the transmitter is operating, overmodulation is found to generate unwanted sideband frequencies which appear for a considerable distance above and below the desired channel. This effect is sometimes called "splatter." These unwanted or "spurious" frequencies, shown in part B of Figure 26-10, cause interference to other stations operating on adjacent channels. It should be clearly understood that overmodulation and its attendant distortion and interference are to be avoided.

In addition to the above problems, over modulation also causes abnormally large voltages and currents to exist at various points within the transmitter. Where sufficient overload protection by circuit breakers and fuses is not provided, these excessive voltages can cause arcing between transformer windings and between the plates of capacitors, permanently destroying the dielectric material. Excessive currents can cause overheating of tubes and other components.

Although it is desirable to operate a transmitter at 100% modulation in order to inject a maximum amount of energy into the sidebands, this ideal condition is seldom possible when a carrier is modulated by voice or music signals. The reason for this is the great and rapid fluctuations in amplitude which these signals normally contain. When the modulator is properly adjusted the loudest parts of the transmission will produce 100% modulation. The quieter portions of the signal then produce lesser degrees of modulation.

In order to measure degrees of modulation less than 100%, it is convenient to use a MODULATION FACTOR (M) to indicate the relative magnitudes of the RF carrier and the audio modulating signal. Numerically, the modulation factor is:

$$M = \frac{E_m}{E_c} \qquad (26\text{-}1)$$

where: M = the modulation factor
E_m = the peak, peak-to-peak, or RMS value of the modulating voltage
E_c = the carrier voltage in the same units as E_m.

To illustrate the use of equation (26-1), assume that a carrier wave having a peak amplitude of 400 volts is modulated by a 3 kc sine wave having a peak amplitude of 200 volts, the modulation factor is:

$$M = \frac{E_m}{E_c} \qquad (26\text{-}1)$$

$$M = \frac{200}{400}$$

$$M = 0.5$$

If the modulation factor is multiplied by 100 the resultant quantity is the PERCENT OF MODULATION (%M). Thus:

$$\% M = \frac{E_m}{E_c} \times 100 \qquad (26\text{-}2)$$

where E_m and E_c have the same meaning as in equation (26-1). When E_m = 200 volts and E_c =

400 volts, the percent of modulation is:

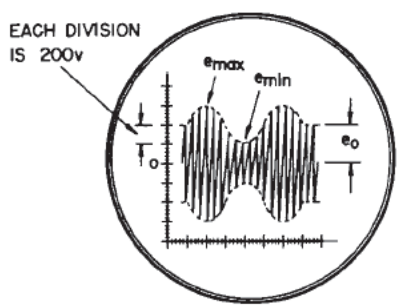
$$\% M = \frac{E_m}{E_c} \times 100 \qquad (26\text{-}2)$$

$$\% M = \frac{200}{400} \times 100$$

$$\% M = 50 \%$$

By using an appropriate equation, the percent of modulation can be determined from the modulation envelope pattern. This method is useful when the percent of modulation is to be determined using the pattern on the screen of an oscilloscope. For example, assume that an oscilloscope connected to the output of a modulator circuit produces the screen pattern shown in Figure 26-11. According to the setting of the calibration control, each large division on the vertical scale is equal to 200 volts. By using this scale, the peak carrier amplitude (unmodulated portion) is seen to be 400 volts. The peak amplitude of the carrier is designated as e_o in Figure 26-11.

The amplitude of the audio modulating voltage can be determined from the amplitude variations in the envelope pattern. Notice, that the peak-to-peak variation in envelope amplitude ($e_{max} - e_{min}$) is equal to 400 volts on the scale. The peak amplitude of the audio voltage is, therefore, 200 volts. If these RF and audio voltage values are inserted into equation (26-2), the



$$\%M = \frac{e_{max} - e_{min}}{2e_o} \times 100$$

—OR—

$$\%M = \frac{e_{max} - e_{min}}{e_{max} + e_{min}} \times 100$$

Figure 26-11 - Computing the percent of modulation from the modulation envelope.

pattern in Figure 26-11 is found to represent 50 % modulation.

If E_m and E_c in equation (26-2) are assumed to represent peak-to-peak values the following derivation results:

$$\% M = \frac{E_m}{E_c} \times 100 \qquad (26\text{-}2)$$

Since the peak-to-peak value of E_m in Figure 26-11 is $e_{max} - e_{min}$:

$$\% M = \frac{e_{max} - e_{min}}{E_c} \times 100$$

and since the peak-to-peak value of the carrier E_c is 2 times e_o, then:

$$\% M = \frac{e_{max} - e_{min}}{2e_o} \times 100 \qquad (26\text{-}3)$$

where e_{max}, e_{min}, and e_o represent the quantities indicated on the pattern in Figure 26-11.

Since, on the screen of a cathode-ray tube, linear vertical distance represents voltage, distance units can be used in place of voltages in equation (26-3). Thus, if only the percent of modulation is required, the oscilloscope need not be calibrated and the actual circuit voltages are not required. In Figure 26-11, e_{max} represents 600 volts or 3 large divisions. Voltage e_{min} is 200 volts or 1 division and e_o is 400 volts or 2 divisions. Using equation (26-3) and the dimensions of the screen pattern:

$$\% M = \frac{e_{max} - e_{min}}{2e_o} \times 100 \qquad (26\text{-}3)$$

$$\% M = \frac{3 - 1}{2 \times 2} \times 100$$

$$\% M = \frac{2}{4} \times 100$$

$$\% M = 50 \%$$

Thus, either scale units or volts can be used to compute the percent of modulation from the screen pattern.

When e_o of equation (26-3) is difficult to measure, an alternate solution can be obtained with equation (26-4) below. The proof of this equation is left as an exercise for the reader.

$$\% M = \frac{e_{max} - e_{min}}{e_{max} + e_{min}} \times 100 \qquad (26\text{-}4)$$

Q3. What would be the percent of modulation if e_{max} and e_{min} had identical values?

26-11. Sideband Voltage and Power

In a previous section it was stated that the intelligence is recovered from an amplitude modulated wave by heterodyning the sideband frequencies with the carrier, and then filtering out all but the audio difference frequency. In order to obtain a large amplitude difference frequency, as much energy as possible should be contained in the sidebands before demodulation. This qualification can be met by maintaining the average percent of modulation as high as practicable without causing overmodulation on peaks.

To gain an understanding of the way in which the percent of modulation affects the sideband energy, it is necessary to examine the relationship between the power contained in the carrier and the power contained in the sidebands. In a transmitter, the carrier and sidebands are applied to a common impedance (tank circuit, antenna, etc.). If the value of this impedance and the voltage across it are known, the power relationships can be computed.

The carrier and modulating voltage were previously found to be related mathematically such that:

$$M = \frac{E_m}{E_c} \qquad (26\text{-}1)$$

Solving this equation for E_m:

$$E_m = M\,E_c \qquad (26\text{-}5)$$

This equation shows that for a fixed carrier amplitude, the voltage that each sideband frequency is capable of producing across a common impedance varies directly with the modulation factor. It must be remembered, however, that the audio modulating voltage E_m is divided between two equal amplitude side frequencies. Thus, the voltage E_{sf} of EITHER side frequency is:

$$E_{sf} = \frac{M\,E_c}{2} \qquad (26\text{-}6)$$

If this side frequency voltage E_{sf} is applied to a load impedance R_L, the power P_{sf} developed by this voltage is:

$$P_{sf} = \frac{(E_{sf})^2}{R_L} \qquad (26\text{-}7)$$

Substitution of equation (26-6) into equation (26-7) yields:

$$P_{sf} = \frac{\left(\frac{M\,E_c}{2}\right)^2}{R_L} \qquad (26\text{-}8)$$

Simplifying equation (26-8)

$$P_{sf} = \frac{M^2\,E_c{}^2}{4\,R_L} \qquad (26\text{-}9)$$

which can be written as:

$$P_{sf} = \frac{M^2}{4} \times \frac{E_c{}^2}{R_L} \qquad (26\text{-}10)$$

The power P_c in the carrier is:

$$P_c = \frac{E_c{}^2}{R_L} \qquad (26\text{-}11)$$

Substituting equation (26-11) for the right-hand quantity in equation (26-10):

$$P_{sf} = \frac{M^2}{4} \times P_c \qquad (26\text{-}12)$$

where:
- M = the modulation factor (using a single frequency, constant amplitude sine wave for modulation)
- P_{sf} = the average power in ONE of the two equal amplitude side frequencies
- P_c = the average carrier power (found using RMS values of E and I)

In the majority of cases the power in one side frequency is of less importance than the TOTAL SIDEBAND POWER P_{sb}. Since two equal amplitude side frequencies exist, the total sideband power is twice the power obtained with equation (26-12). Thus:

$$P_{sb} = 2\frac{M^2}{4} \times P_c \qquad (26\text{-}13)$$

Simplifying:

$$P_{sb} = \frac{M^2}{2} \times P_c \qquad (26\text{-}14)$$

where:
- P_{sb} = the TOTAL sideband power
- M = the modulation factor
- P_c = the average carrier power

A transmitter is normally rated according to the amount of UNMODULATED CARRIER POWER it is designed to deliver to an antenna. If the transmitter is to be used by a commercial broadcast station in the standard broadcast band (535 kc to 1605 kc), the allowable carrier power

A3. Zero percent.

---

is prescribed in the station license issued by the FCC. This unmodulated carrier power is called the station's AUTHORIZED POWER. The sideband power which is generated as a result of modulation exists in addition to the unmodulated carrier power.

For example, assume a 10 kilowatt carrier is modulated 100 %. The sideband power according to equation (26-14) is:

$$P_{sb} = \frac{M^2}{2} \times P_c \qquad (26\text{-}14)$$

Since the percent of modulation is 100, M is equal to one. Substituting values into equation (26-14):

$$P_{sb} = \frac{1^2}{2} \times 10,000$$

$$P_{sb} = 5000 \text{ watts.}$$

Notice, that this computation shows that AT 100 % MODULATION, THE SIDEBANDS CONTAIN ONE-HALF AS MUCH POWER AS THE UNMODULATED CARRIER.

To illustrate the importance of a high percent of modulation, assume that the same 10 kilowatt carrier is modulated only 50 %. The sideband power at 50 % modulation is:

$$P_{sb} = \frac{M^2}{2} \times P_c \qquad (26\text{-}14)$$

Since the percent of modulation is 50, m is equal to 0. 5. Substituting values into equation (26-14):

$$P_{sb} = \frac{0.5^2}{2} \times 10,000$$

$$P_{sb} = \frac{0.25}{2} \times 10,000$$

$$P_{sb} = 1250 \text{ watts}$$

Note, that reducing the percent of modulation to ONE-HALF of its original value, causes the sideband power to decrease to ONE-FOURTH the amount obtained at 100 percent modulation. This shows the importance of a high percent of modulation.

Under certain circumstances it is necessary to compute the total power contained in a modulated wave. The TOTAL POWER P_T is equal to the carrier power plus the sideband power. As an equation:

$$P_T = P_c + P_{sb} \qquad (26\text{-}15)$$

substituting for P_{sb}:

$$P_T = P_c + \frac{M^2}{2} \times P_c \qquad (26\text{-}16)$$

factoring out P_c

$$P_T = P_c \left(1 + \frac{M^2}{2} \right) \qquad (26\text{-}17)$$

To show the application of equation (26-17), the total power radiated from a 50 kilowatt broadcast transmitter will be computed for a condition of 100 percent modulation. Inserting values into equation (26-17):

$$P_T = P_c \left(1 + \frac{M^2}{2} \right) \qquad (26\text{-}17)$$

$$P_T = 50,000 \left(1 + \frac{1^2}{2} \right)$$

$$P_T = 50,000 \, (1.5)$$

$$P_T = 75,000 \text{ watts}$$

Notice, that WHEN MODULATED 100 % THE TOTAL RADIATED POWER IS 1. 5 TIMES THE POWER IN THE UNMODULATED CARRIER.

Q4. What portion of the total radiated power is contained in the carrier at 100 % modulation?

Q5. What portion of the total radiated power is contained in ONE sideband at 60 % modulation?

Q6. What change would occur in a 100 watt carrier, as the percent of modulation is increased from 20 % to 40 % ?

26-12. Modulator Power Requirements

The signal which exists at the output of a plate modulated stage consists of a carrier frequency and two sideband frequencies, each of which represents a finite amount of alternating current power. This power is obtained through a conversion process in which direct current and voltage from the FPA power supply are converted into carrier power, and alternating current and voltage from the modulator are converted into sideband power.

The power distribution for a 400 watt trans-

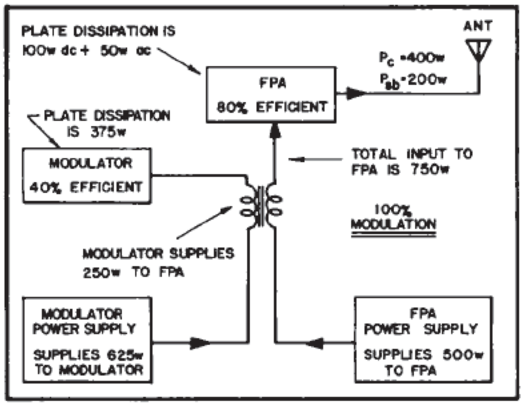


Figure 26-12 - Power distribution in a plate modulated stage.

mitter is shown in Figure 26-12. For this example it will be assumed that the transmitter is modulated 100 percent, and that the final power amplifier and modulator stages have efficiencies of 80 % and 40 % respectively. The plate efficiency of an amplifier can be computed using the following equation:

$$\text{Eff} = \frac{P_o}{P_{in}} \qquad (26\text{-}18)$$

which can be transposed into the following two forms:

$$P_o = P_{in}\,\text{Eff} \qquad (26\text{-}19)$$

and:

$$P_{in} = \frac{P_o}{\text{Eff}} \qquad (26\text{-}20)$$

where: Eff = the efficiency of the amplifier
P_o = the ac output power in watts
P_{in} = the dc input power ($E_b \times I_b$) to the plate of the amplifier

If the modulating signal is reduced to zero, only the carrier frequency will exist in the output of the final power amplifier. Since the final power amplifier operates class C and has an efficiency of 80 %, some power is dissapated at the plate of the tube and the power supply must deliver more than the 400 watts of carrier power that appears at the antenna. The power drawn from the FPA power supply is computed using equation (26-20) as follows:

$$P_{in} = \frac{P_o}{\text{Eff}} \qquad (26\text{-}20)$$

$$P_{in} = \frac{400}{0.80}$$

$$P_{in} = 500 \text{ watts}$$

This computation shows that in order to generate the carrier, 500 watts of dc power must be supplied to the final power amplifier from its power supply. Because the tube is 80 % efficient, 100 of the 500 watts are lost at the plate of the tube as plate dissipation (heat). The remaining 400 watts are converted into carrier power and appear at the antenna.

When the transmitter is 100 percent modulated, two sideband frequencies containing one-half as much power as the carrier appear at the antenna. Since the sidebands occur as a result of modulation, the power which they contain is supplied by the modulator. Numerically, the sideband power at the antenna is 200 watts. This represents 80 % of the audio power applied to the final power amplifier plate circuit by the modulator. Equation (26-20) can be applied to the sidebands to determine the exact amount of power the modulator must supply for 100 % modulation.

$$P_{in} = \frac{P_o}{\text{Eff}} \qquad (26\text{-}20)$$

$$P_{in} = \frac{200}{0.80}$$

$$P_{in} = 250 \text{ watts}$$

Therefore, the modulator must supply 250 watts of audio power to the plate of the final power amplifier inorder to produce 100 % modulation. Of this 250 watts only 200 reach the antenna, with 50 watts being lost at the plate of the final power amplifier as plate dissipation. It should be carefully noted that the same power relationship exists between the dc and audio input powers as exists between the carrier and sideband powers. Namely, that the audio power of 250 watts is exactly one-half of the final dc input power of 500 watts.

Unlike the final, the modulator stage operates with an audio signal and must be operated class A, AB, or B. This severely limits the efficiency that can be obtained in the modulator stage and results in a large power loss at the plate of the modulator. In Figure 26-12 the modulator has an efficiency of 40 %. Because of this low efficiency, only 250 watts of audio power are obtained from the 625 watts of dc power supplied to the plate of the modulator.

4. Two-thirds.

5. Nine percent.

6. No change. The carrier power remains constant regardless of the percent of modulation.

---

odulator plate dissipation accounts for 375 atts of loss. It is most interesting to compare e dc input power to the modulator, to the wer contained in the sidebands. Of the 625 atts of dc input power applied to the modulator ate circuit, only 200 watts are converted into eful sideband energy.

Referring again to the relationship between e dc input power to the final power amplifier d the ac power required from the modulator, e following equation can be used to compute e modulator power required for any percent modulation.

$$P_m = \frac{M^2}{2} \times E_b \times I_b \qquad (26\text{-}21)$$

here: P_m = the output power required from the modulator
M = the modulation factor
E_b = the FPA plate voltage
I_b = the FPA plate current

hus, for 50 % modulation in Figure 26-12:

$$P_m = \frac{(0.50)^2}{2} \times E_b \times I_b$$

nce $E_b \times I_b$ = 500 watts:

$$P_m = \frac{(0.50)^2}{2} \times 500$$

$$P_m = 62.5 \text{ watts}$$

7. How much of this power would become deband power?

VECTOR ANALYSIS

It was stated earlier, without proof, that the odulation envelope results when the instantaeous sums of the carrier and sideband voltes are plotted with respect to time. If such point-by-point summation of these three volt-

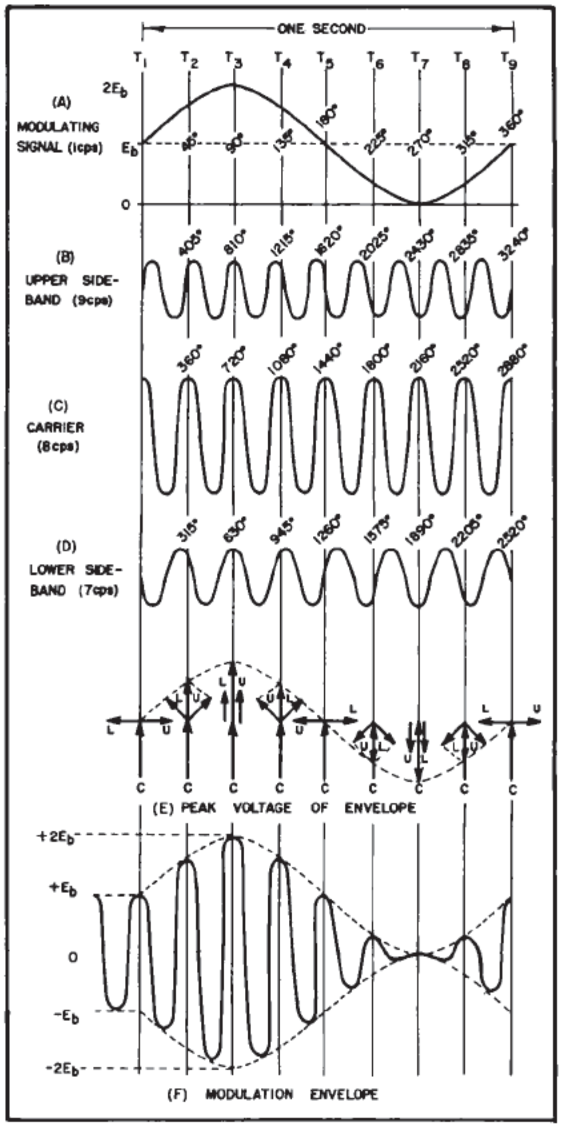


Figure 26-13 - Formation of modulation envelope by addition of vectors representing the carrier and sidebands.

ages were attempted, it would prove to be a monumental task. The same end result can be obtained, however, be using a rotating vector to represent each of the three frequencies in the composite envelope. In the following analysis the vectors will be scaled to indicate the peak value of the frequencies they represent.

To further simplify the analysis, a frequency of 8 cycles per second has been chosen to represent the carrier frequency. Each cycle of the carrier then requires one-eighth of a second to complete 360°. The carrier will be 100 percent

modulated by a sine wave having a frequency of 1 cycle per second, thereby producing sideband frequencies of 7 and 9 cycles per second.

26-13. Envelope Development From Vectors

The modulating signal, upper sideband, carrier, and lower sideband are illustrated in parts A through D respectively in Figure 26-13. Notice, that the vertical lines passing through the diagram divide each waveform into segments of one-eighth second each. These lines also coincide with the starting and ending points of each cycle of the carrier wave.

During the first one-eighth of a second (T_1 to T_2) the carrier wave completes exactly one cycle or 360°. The upper sideband, which has a frequency of 9 cycles per second must complete each cycle in less than one-eighth of a second. Therefore, during the time required for the carrier to complete one cycle of 360°, the upper sideband is able to complete one cycle of 360° plus an additional 45° of the next cycle, for a total of 405°.

The lower sideband has a frequency of 7 cycles per second and thus cannot complete an entire cycle in one-eighth of a second. During the time interval required for the carrier wave to progress through 360°, the lower sideband frequency of 7 cycles per second can complete only 315°, 45° short of a full cycle.

Keeping these facts in mind, it can be seen that the phase angle between the two sideband frequencies, and the phase angle between each sideband frequency and the carrier frequency will continually shift. At one instant of time the carrier and sidebands are in phase, causing the envelope amplitude to be twice the amplitude of the carrier. At another instant of time the sidebands are out of phase with the carrier, causing complete cancellation of RF voltage. The envelope amplitude becomes zero at this point. Thus, although the carrier and sideband frequencies have constant amplitudes, the ever-changing phase difference between them causes the modulation envelope to vary continuously in amplitude.

The vector analysis of the modulation envelope will be developed with the aid of Figure 26-14. In part A of this Figure a vertical vector C has been drawn to represent the carrier wave in Figure 26-13. At time T_1 in Figure 26-13 the upper and lower sideband frequencies are of opposite phase with respect to each other, and 90° out of phase with respect to the carrier. This condition is illustrated in Figure 26-14A by drawing the sideband vectors U and L in opposite directions along the horizontal axis. Since the upper side band U is equal in amplitude but opposite in phase to lower sideband L, the two sideband voltages cancel one another, and

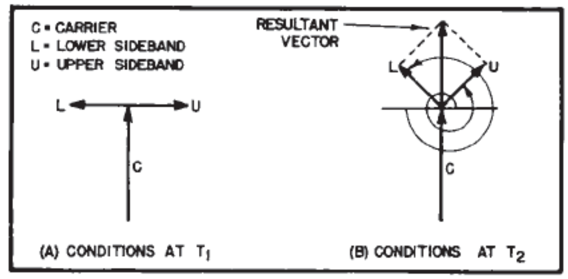


Figure 26-14 - Vector diagrams for time T_1 and T_2.

the amplitude of the envelope at this instant of time is equal to the amplitude of the carrier. This same vector diagram is shown on a smaller scale in Figure 26-13E.

During the one-eighth of a second time interval between T_1 and T_2, all three vectors rotate in a counter-clockwise direction at a velocity determined by their respective frequencies. The vector representing the carrier, for example, has made one complete rotation of 360° and is back in its original position, as shown in Figure 26-14B. The upper sideband frequency, however, must complete 405° in this same one-eighth of a second. Notice in Figure 26-14B, that the vector U has made one complete counter-clockwise rotation of 360° plus an additional 45° for a total rotation of 405°. The vector L representing the lower sideband rotates at a velocity less than that of either the carrier or the upper sideband. In one-eighth of a second vector L completes only 315°, which is 45° short of one complete rotation. Thus, at the end of one-eighth of a second, the three vectors have advanced to the positions shown in Figure 26-14B.

By adding vector U to vector L, the resultant vector in Figure 26-14B is obtained. Since each sideband has one-half the amplitude of the carrier, and the two sidebands differ in phase by 90°, the Pythagorean Theorem can be used to compute the amplitude of the resultant vector. This computation shows the resultant vector to have an amplitude approximately 70 percent that of the carrier. Thus, at time T_2 the amplitude of the modulation envelope is about 1.7 times the amplitude of the carrier. This condition is shown in Figure 26-13F.

By a similar procedure, vector diagrams can be constructed for time intervals T_3 through T_9. This has been done in part E of Figure 26-13. From these nine individual vector diagrams, the complete modulation envelope in Figure 26-13F can be constructed.

Notice in particular, the vector diagrams for T_3 and T_7. At time T_3 all three waves, and therefore all three vectors, are in phase. The modulation envelope at this instant must there-

7. 50 watts

ıre be equal to twice the amplitude of the carıer, since each sideband frequency has one-half ıe amplitude of the carrier.

At T_7 the two sideband frequencies are in ıase with each other but 180° out of phase with ıe carrier. This causes the total sideband ɔltage to cancel the carrier voltage, and the ıodulation envelope becomes zero at this inıtant. Note, that in order for the transmitter ıtput to be zero at T_7, BOTH THE CARRIER ND SIDEBAND FREQUENCIES MUST BE RESENT. If any one of these three frequencies ere missing, complete cancellation would not ccur and RF energy would be present in the ıtput.

Although this vector analysis was made for ·equencies of 7, 8, and 9 cycles per second, ıe same method of attack could be applied to ıe frequencies actually present at the output ſ a transmitter. If this was done the same reılt would be obtained, but with a greater deree of difficulty arising from the large numɜr of cycles per second.

8. What is the amplitude of the modulation ıvelope at time T_5 in Figure 26-13?

5-14. Modulation Level

As stated earlier, the modulating signal can ɜ introduced into any active element of a tube. ı addition to the various arrangements possiıe within a single stage, the modulating signal ın also be applied to any of the RF stages in ıe transmitter. For example, the modulating .gnal could be applied to the control grid or ıate of one of the intermediate power ampliers.

A modulation circuit can usually be placed to one of two categories according to the level ' the carrier wave at the point in the system here the modulation is applied. The FCC denes HIGH LEVEL MODULATION in the Code ' Federal Regulations as: "modulation proıced in the plate circuit of the last radio stage ' the system." This same document defines ƆW LEVEL MODULATION as: "modulation ·oduced in an earlier stage than the final." ɔtice, that a system in which the modulating gnal is applied to the final RF amplifier, but a tube element other than the plate, is left ıdefined.

ADDITIONAL MODULATION SYSTEMS

The plate modulation system is probably the ost frequently used method of amplitude modulation. Its popularity stems from the fact that, compared to other methods, plate modulation has high plate circuit efficiency, good linearity (low distortion), and is simple to adjust. Its only disadvantage is the large amount of audio power required from the modulator.

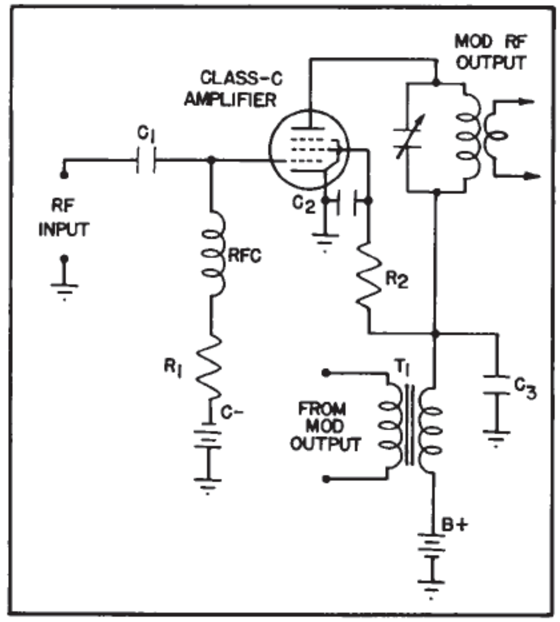


Figure 26-15 - Circuit showing combination plate and screen grid modulation.

26-15. Screen Grid Tubes

When triode tubes are used in the final power amplifier stage, carefully adjusted neutralization circuits must be employed. These circuits, and their critical adjustments, can be avoided by using screen grid tubes in place of triodes for the final power amplifier stage.

As pointed out earlier, the plate current of a screen grid tube is almost independent of plate voltage. The plate current is, however, very much dependent on screen grid voltage. Thus, the modulating voltage is normally applied to the screen and plate circuit simultaneously when plate modulating a screen grid tube.

A combination plate and screen grid modulated amplifier is shown in Figure 26-15. Notice, that the audio modulating voltage across the secondary of the modulation transformer T_1 is in series with the dc supply voltage to both the plate and screen circuits.

In order to prevent screen degeneration at the carrier frequency, capacitor C_2 is connected between the screen grid and cathode of the tube. The value of this capacitor is chosen so that its reactance approaches a short-circuit at

radio frequencies, but appears as an open circuit to the audio modulating frequencies.

Capacitor C_3 serves a purpose similar to that of capacitor C_2. This capacitor must prevent RF frequencies from developing across the secondary of T_1 while having little if any effect on the audio voltages present across the secondary of T_1.

26-16. Control Grid Modulation

A disadvantage of plate modulation is the large amount of power that the modulator must supply to the plate of the RF amplifier (one-half the dc input power to the final for 100% modulation). By applying the modulating voltage to the grid circuit of the RF amplifier, the modulator power requirements can be reduced considerably.

A circuit for producing control grid modulation is shown in Figure 26-16. With the exception of the control grid circuit, the stage resembles a conventional triode class C amplifier stage. Capacitor C_n is the neutralizing capacitor and is used to balance out the undesirable effects of feedback which occurs through C_{gp} and stray circuit capacitances. By employing a tetrode or pentode tube in the RF stage the neutralization circuits can normally be omitted.

The grid circuit in Figure 26-16 resembles that of a normal class C amplifier except for the inclusion of modulation transformer T_1 in series with the grid bias supply E_{cc}. Choke coil RFC and capacitor C_2 are included to prevent the RF energy of the carrier from entering modulation transformer T_1 or bias supply E_{cc}.

During periods of 100 percent modulation, the most positive peaks of the grid signal will drive the grid positive and draw grid current. This places a heavier load on both the modulator and the grid bias supply during the time

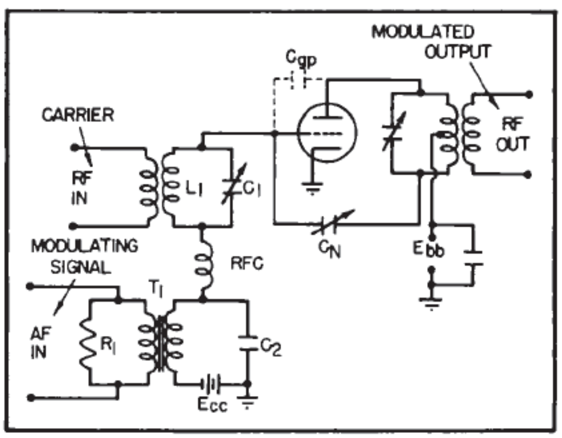


Figure 26-16 - Circuit for producing grid modulation.

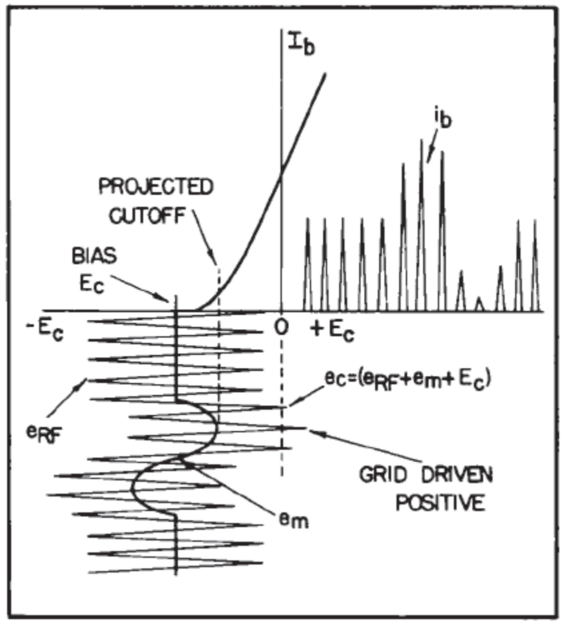


Figure 26-17 - Curve showing operating characteristics of grid modulated amplifier.

interval that grid current is drawn. To prevent fluctuations in bias as a result of grid current the bias supply must be well regulated.

To minimize the loading effect that grid current has on the modulator, the modulator stage should have a low internal impedance. This can be accomplished by using negative feedback loops in the modulator amplifier circuits. Additional stability can be gained by connecting a resistor (R_1) across the primary of the modulation transformer. This resistor placed a constant load on the modulator and thus makes the grid current a small portion of the total load on the modulator.

The operation of a grid modulator circuit can be better understood by referring to Figure 26-17. This illustration shows the dynamic transfer characteristic for the modulated RF amplifier. The bias voltage for the amplifier is adjusted such that the characteristics of the circuit are essentially those of a class C amplifier.

During the application of an audio modulating voltage (e_m), the audio voltage has the effect of varying the bias above and below the average value of bias (E_c). This causes the operating point for the RF carrier to be shifted first to the right, then to the left along the E_c axis, in accordance with the modulating signal. As a result of the shift in operating point, the amplitude of the plate current pulses change in step with the amplitude variations of the modulating signal. When applied to the plate tank

A8. The same amplitude as the unmodulated carrier.

---

circuit, the plate current pulses produce the normal modulation envelope.

The diagram in Figure 26-17 shows that the positive peaks of the carrier wave must be able to shift upward on the curve during the positive alternation of the modulating signal. To allow for this, the amplitude of the carrier applied to the grid is smaller than would be the case for a plate modulated amplifier. During periods of time when the carrier is unmodulated the peak amplitude of the plate current pulses is only about one-half the amplitude they have at the peak of the modulating signal. This means that the carrier power obtainable is considerably less than the amount the tube can produce as a conventional class C amplifier. The small amplitude of carrier input also results in decreased efficiency.

Compared to plate modulation, grid modulation is less efficient, produces more distortion, and requires the final power amplifier to supply all of the power in the output signal. Grid modulation has the advantage of not requiring much power from the modulator.

26-17. Suppressor Grid Modulation

A circuit in which the modulating signal is applied to the suppressor grid is shown in Figure 26-18. In this circuit the suppressor grid is always maintained negative by the suppressor grid bias supply and, therefore, the modulator does not have to supply any power to the modulated amplifier.

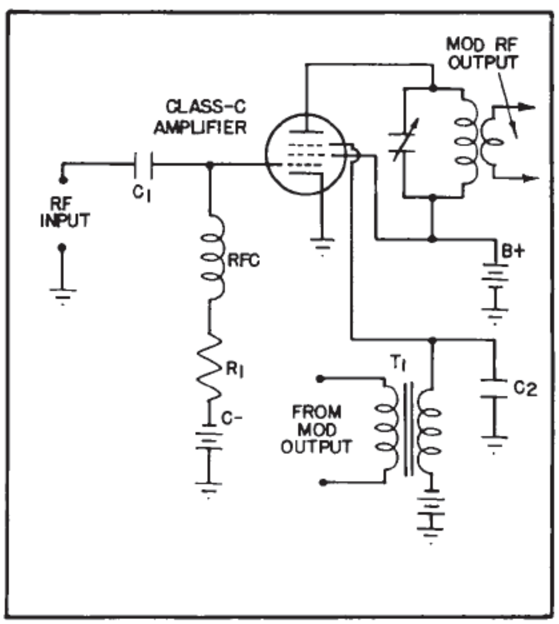


Figure 26-18 - Circuit for suppressor grid modulation.

26-18. Cathode Modulation

A circuit using cathode modulation is shown in Figure 26-19. Since the cathode element is common to both the control grid and plate circuits of the tube, a cathode modulation circuit has characteristics midway between those of plate and control grid modulation. Like suppressor grid modulation, cathode modulation is seldom encountered and therefore will not be elaborated on further.

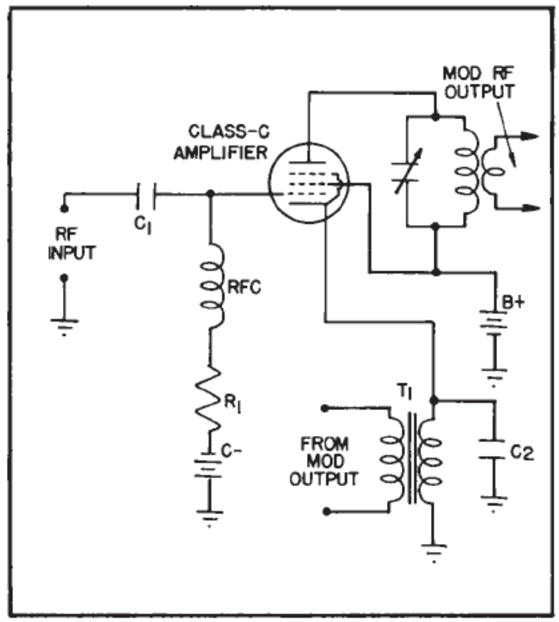


Figure 26-19 - Cathode modulation circuit.

26-19. Checking The Percentage Of Modulation

A transmitter operating at more or less than 100 percent modulation is operating with reduced efficiency. Less than 100 percent modulation causes reduced side band power and reduced transmitting range. Overmodulation results in distortion of the transmitted signal and generation of additional frequencies beyond the stations assigned bandwidth, with possible interference to the operation of adjacent stations. It is therefore desirable that the station have some method of continuously checking or monitoring the percentage of modulation of the transmitted signal. In some cases a meter is used as an indicator and in other cases an oscilloscope is used.

Two methods of connecting an oscilloscope in order to monitor the percentage of modulation are shown in Figure 26-19.

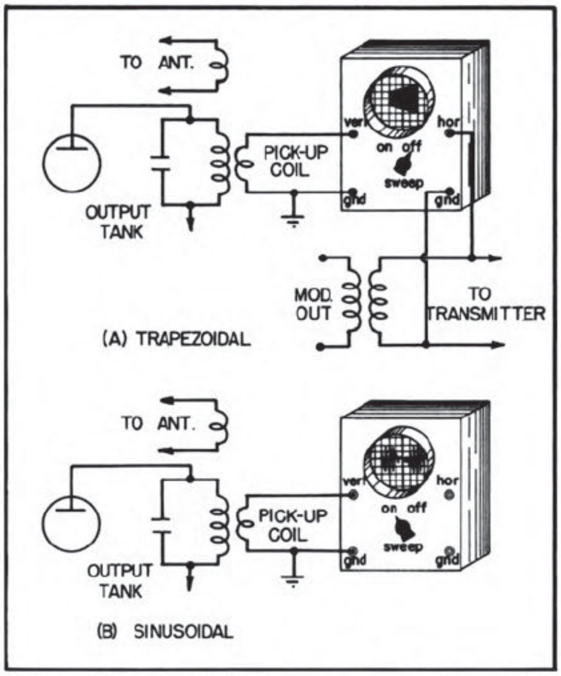


Figure 26-20 - Scope connections for monitoring modulation.

The TRAPEZOIDAL method, shown in part A of the Figure, requires that the transmitted signal be connected to the vertical input of the scope and the modulation signal be connected to the horizontal input of the scope. Due to the high power and voltages present in the output tank it is extremely impractical to connect the scope directly across the tank. For this reason the vertical input is connected to a turn or two of wire, called a pick up coil, which by inductive coupling applies a portion of the transmitted signal to the scope. In part A of the Figure the modulating signal is shown connected directly to the horizontal input. This is permissable PROVIDING the voltages present DO NOT EXCEED the input ratings of the scope. If the voltages present are too large, then a resistive voltage divider can be used so that only a portion of the modulating signal is applied to the horizontal input. Notice that use of the modulating signal to supply the horizontal sweep (time base) requires that the internal sweep of the scope be turned off (in many scopes the sweep off position is labeled HOR. IN, EXT. SWEEP, etc.).

Part B of Figure 26-20 shows the SINUSOIDAL method of monitoring the percentage of modulation. This method only requires that a portion of the transmitted signal be applied to the vertical input of the scope. Notice that the internal sweep of the scope is used.

Figure 26-21 illustrates a comparison between trapezoidal and sinusoidal scope patterns for various percentages of modulation. Part A of the Figure shows the scope displays for zero modulation. In other words, with no modulating signal and only the carrier present. Parts B, C, and D show 50%, 100%, and greater than 100% respectively.

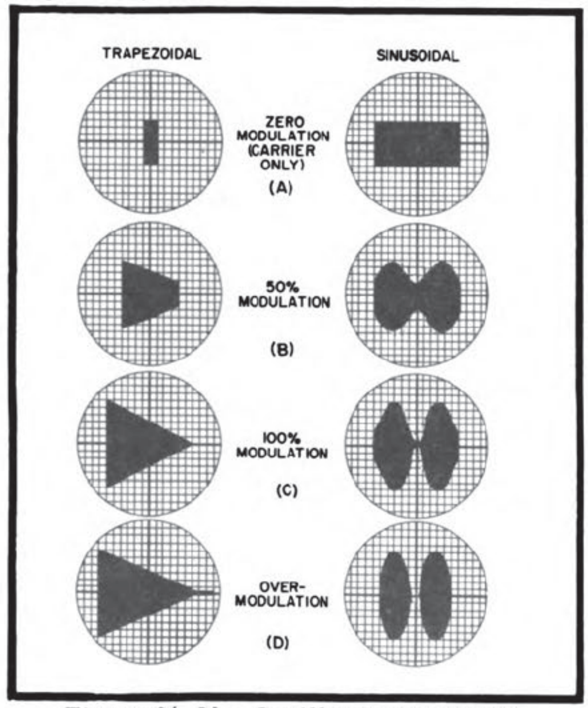


Figure 26-21 - Oscilloscope patterns.

EXERCISE 26

1. What is CW transmission.
2. What is the meaning of the word "modulation?"
3. What is "amplitude modulation?"
4. Explain the phenomenon of heterodyne action.
5. What type of impedance is necessary in order for heterodyning to occur?
6. What frequencies are present across a nonlinear impedance to which signals of 700 kc, 3 kc, and 7 kc are applied?
7. What are side frequencies?
8. What are sidebands?
9. How many side frequencies exist for each modulating frequency?
10. What determines the bandwidth of an AM wave?
11. Where does the intelligence actually exist in an AM wave?
12. Describe the operation of a simple plate modulator circuit.
13. What is meant by the "percent of modulation" of an AM wave?
14. At 100% plate modulation, what relationship exists between the FPA dc supply voltage and the peak RF voltage across the plate tank?
15. At 100% modulation, how does the voltage in one sideband frequency compare with the carrier voltage?
16. At 100% plate modulation what relationship exists between the dc plate supply voltage and the peak amplitude of one side frequency?
17. What percent of modulation exists when the peak RF voltage is 4000 volts, and the peak audio modulating voltage is 2400 volts?
18. What peak value of audio voltage would be necessary to modulate a carrier wave of 2500 volts peak to 60%?
19. Describe overmodulation and its effects.
20. How does the total sideband power compare to the carrier power at 100% modulation?
21. Describe the effects that would occur in a 100% plate modulated class C amplifier if the RF drive to the grid is reduced.
22. What is the total radiated power if the power in one sideband is 50 watts at 100% modulation.
23. If the carrier power is 200 watts at 20% modulation, what would it be at 40% modulation?
24. What power is required from a plate modulator, if a FPA stage having a dc input power of 2000 watts is to be modulated 100%?
25. What is the FPA plate dissipation at 100% modulation, if the carrier power is 800 watts?
26. Explain the meaning of "high level modulation."
27. Explain the meaning of "low level modulation."
28. Describe the operation of a grid modulated stage.
29. Why is a high percent of modulation desirable?
30. Describe the use of the wave envelope and trapezoidal oscilloscope patterns to determine the percent of modulation.

CHAPTER 27

TRANSMISSION LINES

The output of a communications transmitter is a radio frequency signal, the power of which may be low, intermediate, or high. The output frequency may range from a few kilocycles to thousands of megacycles.

The output of the transmitter must be transferred to the load, which is usually an antenna. A device known as a transmission line is used to conduct or guide the energy from the transmitter to the load. A transmission line must be capable of handling the power output of the transmitter, and possess low effective resistance to keep power losses at a minimum.

TYPES OF TRANSMISSION LINES

There are five types of transmission lines that will be discussed in this chapter. They are: the parallel two wire, the twisted pair, the shielded pair, the concentric coaxial line (rigid and flexible), and waveguides. The use of a particular line depends, among other things, on the applied frequency, the power handling capabilities, and the type of installation.

27-1. Two Wire Open Line

One type of parallel line is the two wire open line illustrated in Figure 27-1. This line consists of two wires that are generally spaced from two to six inches apart. This type of line is most often used for power lines, rural telephone lines, and telegraph lines. It is sometimes used as a transmission line between antenna and transmitter or antenna and receiver. An advantage of this type of line is its simple construction. The principal disadvantages of this type of line are the high radiation losses and noise pickup due to the lack of shielding. Radiation losses are produced by the changing fields which are produced by the changing current in each conductor. Some of these lines of force will be radiated from the transmission line in much the same manner as energy is radiated from the sun.

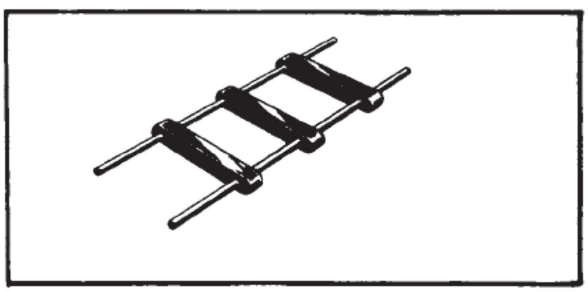


Figure 27-1 - Parallel two-wire line.

Another type of parallel line is the twin lead or two wire ribbon type. This line is illustrated in Figure 27-2. This line is essentially the same as the two wire open line, except that uniform spacing is assured by imbedding the two wires in a low loss dielectric usually polyethylene. Since the wires are imbedded in a thin ribbon of polyethylene, the dielectric space is partly air and partly polyethylene.

27-2. Twisted Pair

The twisted pair transmission line is illustrated in Figure 27-3. As the name implies, the line consists of two insulated wires, twisted to form a flexible line without the use of spacers. It is not used for high frequencies due to the high

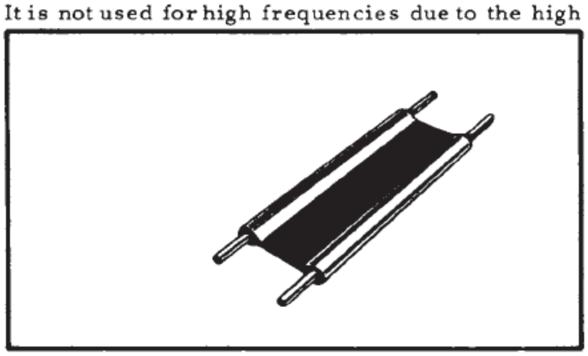


Figure 27-2 - Two-wire ribbon type lines.

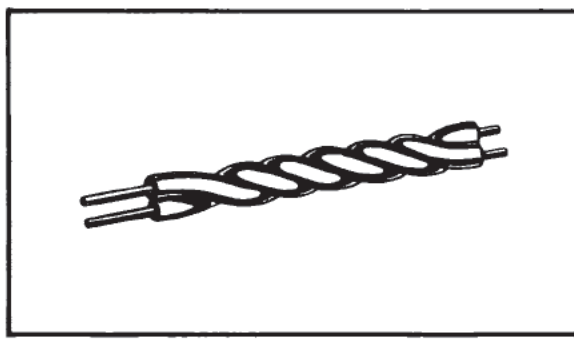


Figure 27-3 - Twisted pair.

losses that occur in the rubber insulation. When the line is wet, the losses increase greatly.

27-3. Shielded Pair

The shielded pair, shown in Figure 27-4, consists of parallel conductors separated from each other, and surrounded by, a solid dielectric. The conductors are contained within a copper braid tubing that acts as a shield. The assembly is covered with a rubber or flexible composition coating to protect the line from moisture or mechanical damage. Outwardly, it looks much like the power cord of a washing machine or refrigerator.

The principal advantage of the shielded pair is that the conductors are balanced to ground; that is, the capacitance between the cables is uniform throughout the length of the line. This balance is due to the grounded shield that surrounds the conductors with a uniform spacing along their entire length. The copper braid shield isolates the conductors from stray magnetic fields.

27-4. Coaxial Lines

There are two types of coaxial lines: the rigid or air coaxial line and the flexible or solid coaxial line. The physical construction of both types is basically the same, each contains two concentric conductors.

The rigid air coaxial line consists of a wire mounted inside of, and coaxially with, a tubular outer conductor. This line is shown in Figure 27-5. In some applications the inner conductor is also tubular. The inner conductor is insulated from the outer conductor by insulating spacers, or beads, at regular intervals. The spacers are made of pyrex, polystyrene, or some other material possessing good insulating characteristics and low loss at high frequencies.

The chief advantage of this type of line is its ability to minimize radiation losses. The electric and magnetic fields in the two wire parallel line extend into space for relatively great distances and radiation losses occur. No electric or magnetic fields extend outside of the outer conductor in a coaxial line. The fields are confined to the space between the two conductors,

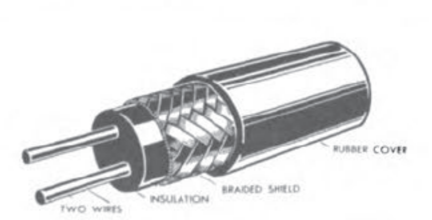


Figure 27-4 - Shielded pair.

Figure 27-5 - Air coaxial.

thus the coaxial line is a perfectly shielded line. Noise pickup from other lines is also prevented.

This line has several disadvantages: it is expensive to construct, it must be kept dry to prevent excessive leakage between the two conductors, and although high frequency losses are somewhat less than in previously mentioned lines they are still excessive enough to limit the practical length of the line.

The condensation of moisture is prevented in some applications by the use of an inert gas, such as nitrogen, helium, or argon, pumped into the line at a pressure of from 3 to 35 pounds per square inch. The inert gas is used to dry the line when it is first installed and a pressure is maintained to insure that no moisture enters the line.

Concentric cables are also made with the inner conductor consisting of flexible wire insulated from the outer conductor by a solid, continuous insulating material. Flexibility may be gained if the outer conductor is made of metal braid. Early attempts at obtaining flexibility employed the use of rubber insulators between the two conductors. The use of rubber insulators caused excessive losses at high frequencies and the bead arrangement allowed moisture carrying air to enter the line, resulting in high leakage current and arc over when high voltages were applied. These problems were solved by the development of polyethylene plastic, a solid substance that remains flexible over a wide range of temperatures. Polyethylene is unaffected by sea water, gasoline, oils and other liquids that may be found aboard ship. High frequency losses due to the use of polyethylene, although greater than the losses would be if air

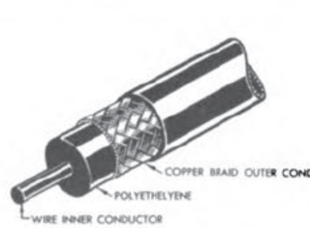


Figure 27-6 - Flexible coaxial.

was used, are lower than the losses resulting from the use of most other practical solid dielectric material. Solid flexible coaxial transmission lines are frequently used by the Navy because of their flexibility, (Figure 27-6).

27-5. Waveguides

The waveguide is classified as a transmission line. However, the method by which it transmits energy down its length differs from the conventional. The operation of the waveguide will not be considered in this chapter because the discussion involves a fundamental knowledge of electromagnetic theory that will be subsequently discussed.

The common types of waveguides are the cylindrical and the more frequently used rectangular type. These waveguides are illustrated in Figure 27-7. The term waveguide is applicable to all types of transmission lines in the sense that they are all used to direct or guide energy from one point to another, and as such it does not matter whether the line is composed of a single conductor, two or more conductors, a coaxial line, a hollow metal tube, or a dielectric rod. Usage, however, has limited the meaning of the word to the hollow metal tube and the dielectric transmission line. The term waveguide, as used in this chapter, means hollow metal tube.

The transmission of electromagnetic energy along a waveguide is closely related to the transmission of electromagnetic energy through space. The space of travel of electromagnetic energy through a waveguide is slower than the transmission of energy through space.

Waveguides may be classed according to cross section (rectangular, elliptical, or circular), or according to material (metallic or dielectric). Dielectric waveguides are seldom used because the losses for all known dielectric materials are too great to efficiently transfer the electric and magnetic fields.

The installation of a waveguide transmission system is somewhat more difficult than the installation of other types of lines. The radius of bends in the line must be greater than two wavelengths, to the operating frequency to avoid excessive attenuation. The cross section must remain uniform around the bend. These difficulties hamper installation in restricted spaces. If the guide is dented, or if solder is permitted to run inside the joints, the attenuation of the line is greatly increased. Dents and obstructions in the guide also reduce the breakdown voltage of the guide. Such faults limit the power handling capability of the system and increase the possibility of arc over. Unless great care is exercised in the installation, one or two carelessly made joints could nullify the initial advantage obtained from the use of the waveguide.

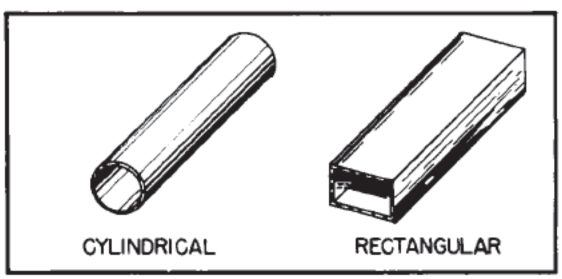


Figure 27-7 - Hollow waveguides. (a) cylindrical (b) rectangular.

Waveguide theory will be covered completely in a later chapter.

27-6. Electrical Characteristics of Transmission Lines

The end of a two wire transmission line that is connected to a source is ordinarily called the GENERATOR END or the INPUT END. The other end of the line, if connected to a load, is called the LOAD END or the RECEIVING END.

The electrical characteristics of the two wire transmission lines are dependent primarily on the construction of the line. Since the two wire line is nothing more than a long capacitor the change of its capacitive reactance will be noticeable as the frequency applied to it is changed. Since the long conductors will have a magnetic field about them when electrical energy is being passed through them, the properties of inductance will also be observed. The values of the inductance and capacitance present are dependent on various physical factors, all of which have been previously discussed. The effects of the line's associated reactances will be dependent on the frequency applied. Since no dielectric is perfect, electrons will manage to move from one conductor to the other through the dielectric; there will be a conductance value for each type of two wire transmission line. This conductance value will represent the value of current flow that may be expected through the insulation. If the line is uniform (all values equal at each unit length) one small section of the line may be represented as shown in Figure 27-8. Such a diagram may represent several feet of line.

The values of resistance, capacitance, inductance, and conductance in the circuit in Figure 27-8, if given, would represent LUMPED values. It is convenient to analyze one section of the line at a time. To adequately do this, the values should be significant. This type of analysis is valid because at some operating frequency, these values will become highly significant, and an integral part of either the

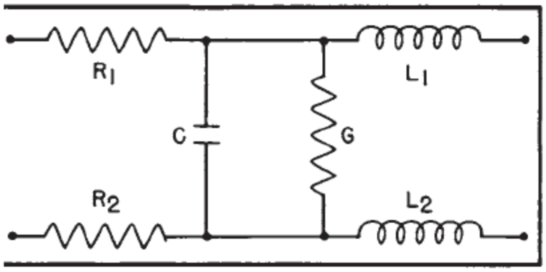


Figure 27-8 - Equivalent circuit of a two-wire transmission line.

ɪrcuit symbols for Figure 27-8:

- L_1 = inductance of top wire
- L_2 = inductance of bottom wire
- R_1 = resistance of top wire
- R_2 = resistance of bottom wire
- G = conductance between wires
- C = capacitance between wires

ɪtput circuit of a stage, or part of the input rcuit of the next stage. These values can ɪuse decreased gain, distortion, and undesirɪle phase shifts. To prevent these undesirable fects, the lumped values may be incorporated to the circuit as usable components, or comnsated for, or simply tolerated. However, it important to know their effects to be able understand the circuits that employ these lues as usable components.

In many applications, the values of conductɪce and resistance are very small and may be glected. If they are neglected, the circuit will ɪpear as shown in Figure 27-9. Notice that is network is TERMINATED with a resistance at represents the impedance of an infinite ɪmber of sections exactly like the section of ɪe under consideration. The termination is ɪnsidered to be a load connected to the line.

A line infinitely long would be composed of ɪ infinite number of inductors and capacitors. a voltage is applied to the input terminals of e line, current would begin to flow. It would

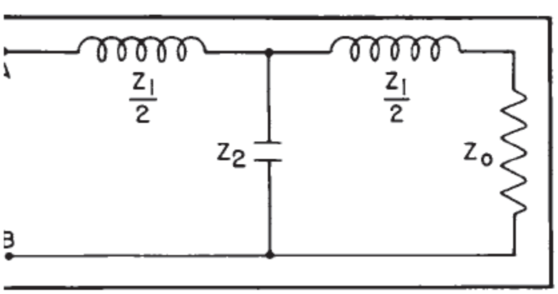


Figure 27-9 - Simplified circuit terminated with its characteristic impedance.

continue to flow as long as the capacitors and inductors were able to assume a charge. Since there are an infinite number of these sections of line, the current would flow indefinitely. If the infinite line is uniform, the impedance of each section would be the same as the impedance offered to the circuit by any other section of line of the same unit length. Therefore, the current would be of some finite value. If the current flowing in the line, and the voltage applied across it are known, the impedance of the infinite line could be determined by using Ohm's law. This impedance is called the CHARACTERISTIC IMPEDANCE of the line. The symbol used to represent the characteristic impedance is Z_o. If by some means, the characteristic impedance of the line was measured at any point on the line, it would be found that the characteristic impedance would be the same. The characteristic impedance is also called the SURGE IMPEDANCE.

If Figure 27-9, the distributed inductance of the line is divided equally into two parts in the horizontal arms of the T. The distributed capacitance is lumped and shown connected in the central leg of the T. The line is terminated in a resistance equal to that of the characteristic impedance of the line as seen from terminals A and B. The reasons for using this value of resistive termination will be fully explained in section 27-8 of this chapter. Since the circuit in Figure 27-9 is nothing more than a series-parallel LCR circuit, the impedance of the network may be determined by the formula that will now be developed.

By the product over sum method:

$$Z_o = \frac{Z_1}{2} + \frac{Z_2\left(\frac{Z_1}{2} + Z_o\right)}{Z_2 + \frac{Z_1}{2} + Z_o}$$

simplifying:

$$Z_o = \frac{Z_1}{2} + \frac{\frac{Z_1 Z_2}{2} + Z_o Z_2}{Z_2 + \frac{Z_1}{2} + Z_o}$$

Expressing the right-hand member in terms of the least common denominator:

$$Z_o = \frac{Z_1 Z_2 + \frac{Z_1^2}{2} + Z_1 Z_o + \frac{2Z_1 Z_2}{2} + 2Z_o Z_2}{2\left(Z_2 + \frac{Z_1}{2} + Z_o\right)}$$

If both sides of this equation are multiplied by the denominator of the right hand member, the result is:

$$2Z_2Z_o + \frac{2Z_1Z_o}{2} + 2Z_o^2 =$$

$$Z_1Z_2 + \frac{Z_1^2}{2} + Z_1Z_o + \frac{2Z_1Z_2}{2} + 2Z_oZ_2$$

Simplifying:

$$2Z_o^2 = 2Z_1Z_2 + \frac{Z_1^2}{2}$$

or:

$$Z_o^2 = Z_1Z_2 + \left(\frac{Z_1}{2}\right)^2$$

If the transmission line is to be accurately represented by an equivalent network, the T-network section of Figure 27-9 must be replaced by an infinite number of similar sections. Thus, the distributed inductance in the line will be divided into n sections, instead of the number (2) as indicated in the last term of the preceding equation. As the number of sections approaches infinity, the last term $\frac{Z_1}{n}$ will approach zero.

Therefore: $$Z_o = \sqrt{Z_1Z_2}$$

Since the term Z_1 represents the inductive reactance, and the term Z_2 represents the capacitive reactance:

$$Z_o = \sqrt{2\pi f\ L \times \frac{1}{2\pi f\ C}}$$

and:

$$Z_o = \sqrt{\frac{L}{C}} \qquad (27\text{-}1)$$

The last formula indicates that the characteristic impedance of the line depends on the ratio of the distributed inductance and capacitance in the line. An increase in the separation of the wires increases the inductance and decreases the capacitance. This effect takes place because the effective inductance is proportional to the flux which may be established between the two wires. If the two wires carrying current in opposite directions are placed farther apart, more magnetic flux is included between them (they cannot cancel their magnetic effects as completely as if the wires were closer together), and the distributed inductance is increased. The capacitance is lowered if the plates of the capacitor, (in this case the plates are the two wires), are more widely spaced.

Thus, the effect of increasing the spacing of the two wires is to increase the characteristic impedance, because the L/C ratio is increased. Similarly, a reduction in the diameter of the wires also increases the characteristic impedance. The reduction in the size of the wire affects the capacitance more than the inductance, for the effect is equivalent to decreasing the size of the plates of a capacitor in order to decrease the capacitance. Any change in the dielectric material between the two wires also changes the characteristic impedance. If a change in the dielectric material increases the capacitance between the wires, the characteristic impedance, by equation 27-1, is reduced.

The characteristic impedance of a two wire line with air as the dielectric may be obtained from the formula:

$$Z_o = 276 \log_{10} \frac{2D}{d} \qquad (27\text{-}2)$$

where D is the spacing between the wires (center to center), and d is the diameter of one of the conductors.

The characteristic impedance of a concentric or coaxial line, with an air dielectric, also varies with L and C. However, because the difference in construction of the two lines causes L and C to vary in a slightly different manner, the following formula must be used to determine the characteristic impedance of the coaxial line:

$$Z_o = 138 \log_{10} \frac{D}{d} \qquad (27\text{-}3)$$

where D is the inner diameter of the outer conductor and d is the outer diameter of the inner conductor.

Q1. Explain what happens to the characteristic impedance of a two-wire transmission line when the separation between the conductors is increased.

27-7. Terminating the Line

If a line is terminated with its characteristic impedance, the same value of current would flow in the circuit as would be realized in an unterminated, infinitely long line. Since the value of the terminating impedance is made equal to the impedance of the generator, and if the impedance ordinarily matches the impedance

A1. The impedance increases because the capacitance decreases as the distance between the conductors increases. The L/C ratio increases resulting in an increased characteristic impedance.

of the line; there would exist a maximum transfer of power from the generator to the line, and from the line to its terminating impedance. It may be said that under these conditions, all of the power sent down the line will be absorbed by the terminating impedance. This, of course is only true of a loss free line.

There are other methods of line termination. The line may be terminated in either an open or a short or have an R_L not equal to Z_o. The characteristics of a line when it is terminated in any of the above conditions will be discussed later.

27-8. Propagation of DC Voltage Down a Line

To better understand the characteristics of a transmission line with an ac voltage applied, the infinitely long transmission line will first be analyzed with a dc voltage applied. This will be accomplished using the circuit illustrated in Figure 27-10. In this circuit the resistance of the line is not shown. The line will be assumed to be loss free.

Considering only the capacitor C_1 and the inductor L_1 as a series circuit, when voltage is applied to the network through the ammeter; capacitor C_1 will have the ability to charge through inductor L_1. It is characteristic of an inductor that, at the first instant of time when voltage is applied, a maximum voltage is developed across it and minimum current is permitted to pass through it. At the same time, the capacitor will have a minimum of voltage across it and a tendency to pass a maximum current. The maximum current is not permitted to flow at the first instant because of the action of the inductor which is in the charge path of the capacitor. The voltage across points c and d, at this time is zero. Since the remaining portion of the line is connected to points c and d, there will be zero volts developed across the line at the first instant of time. The voltage across the line is dependent on the charging action of the capacitor, C_1. It will require some finite amount of time for capacitor C_1 to charge through inductor L_1. As capacitor C_1 is charging, the ammeter records the charging current. When C_1 charges to a voltage which is near the value of the applied voltage, capacitor C_2 will begin to charge through inductors L_1 and L_2. The charging of capacitor C_2 will again require time. In fact the time required for the voltage to reach points e and f will be the same time as was necessary for the original voltage to reach points c and d. This is true because the line is uniform, and the values of the reactive components are the same throughout its entire length. This action will continue in the same manner until all of the capacitors in the line are charged. Since the number of capacitors in an infinite line is infinite, the time required to charge the entire line would be an infinite amount of time. It is important to note that current is flowing continuously in the line, and that it has some finite value. A circuit that will display current characteristics similar to the charging of an infinitely long line is shown in Figure 27-11.

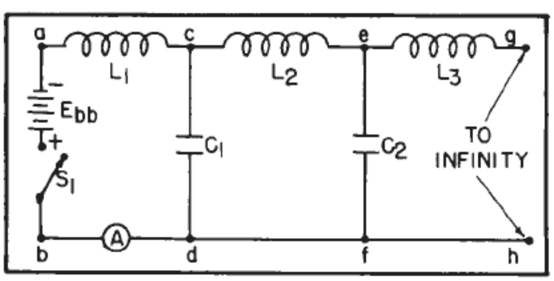


Figure 27-10 - DC voltage applied to a transmission line.

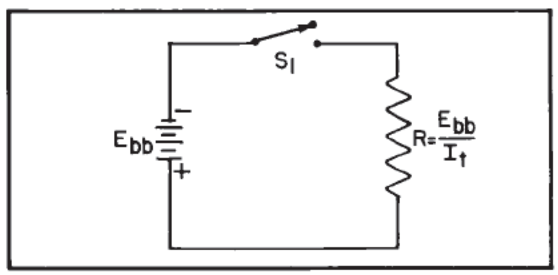


Figure 27-11 - Equivalent circuit.

If the voltage applied to the line and the resulting current through it were known, a resistance value using Ohm's law could be computed. The resistance value thus found is often called the CHARACTERISTIC RESISTANCE of the line. However, the term most often applied is "characteristic impedance". If along the line, the circuit was broken, and the battery connected across the resulting new terminals; the value of the resulting current would be the same. This circuit is shown in Figure 27-12.

In Figure 27-12A, the line with the batteries connected to the normal generator end is shown. The ammeter will indicate the value of current flowing. If the line is broken at points c and d, and the source connected as shown in Figure

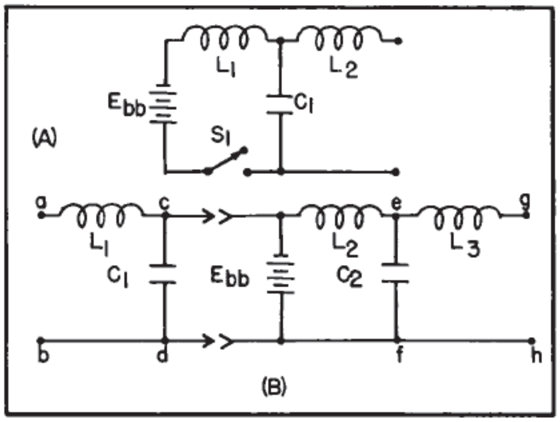


Figure 27-12 - Changing the physical location of the source.

27-12B; the resulting current measured on the ammeter would be the same as it was in the original circuit (Figure 27-12A). This experiment gives further proof to the fact that the impedance of a uniform infinite line is the same throughout its length.

If the battery is re-connected to the input terminals, a and b, and a resistance equal to the characteristic impedance of the line is connected between points c and d, the current recorded by the ammeter would again be the same. The transmission lines considered thus far have been loss free lines. If a practical transmission line, one possessing resistance and shunt conductance (Figure 27-8), were used; power would be dissipated by the line in the form of heat.

When an electrostatic field is moving down the line, its associated electric and magnetic fields are said to be PROPAGATED down the line. It was found and mentioned previously that time was required to charge each unit section of the line, and if that line was infinitely long; the line would require an infinitely long time to charge. The time for a field to be propagated from one point on a line to another may be computed, for if the time and the length of the line are known, the VELOCITY OF PROPAGATION may be determined. The network shown in Figure 27-13 is the circuit that will be used to compute the time required for the voltage wavefront to pass a section of line of specified length. The total charge in coulombs on capacitor C_1 is determined by the relationship:

$$Q = CE \qquad (10\text{-}11)$$

Since the charge on the capacitor in the line had its source at the battery, the total amount of

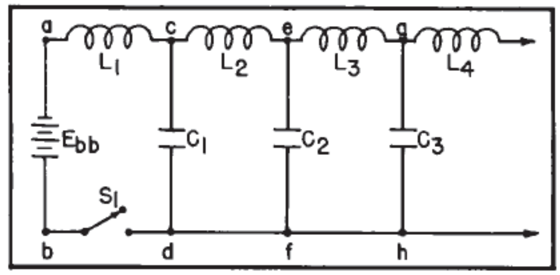


Figure 27-13 - Circuit for computing time of travel.

charge removed from the battery will be equal to:

$$Q = IT \qquad (5\text{-}1)$$

Because these charges are equal, they may be equated:

$$CE = IT$$

As the capacitor C_1 charges to (CE), capacitor C_2 contains a zero charge. Since capacitor C_1 voltage is distributed across C_2 and L_2, at the same time the charge on C_2 is practically zero, the voltage across C_1 (points c and d) must be, by Kirchhoff's law, entirely across L_2. The value of the voltage across the inductor is given by the equation:

$$e = L\frac{di}{dt} \qquad (9\text{-}3)$$

Since current and time start at zero, the change in time and the change in current are equal to the final current and the final time. Equation 9-3 becomes:

$$ET = LI$$

Solving the equation for I:

$$I = \frac{ET}{L}$$

Solving the equation that was a statement of the equivalency of the charges for current:

$$I = \frac{CE}{T}$$

Equating both of these expressions:

$$\frac{ET}{L} = \frac{CE}{T}$$

Solving the equation for T:

$$T^2 = LC$$

or:

$$T = \sqrt{LC} \qquad (27\text{-}4)$$

Since velocity is a function of both time and distance, the formula for computing propagation velocity is:

$$V_p = \frac{D}{\sqrt{LC}} \qquad (27\text{-}5)$$

where: V_p = velocity of propagation
D = distance of travel
$\sqrt{LC}$ = time

It should again be noted that the time required for a wave to traverse a transmission line segment will depend on the values of L and C, these values will be different, depending on the type of transmission line considered.

Q2. If the capacitance of a line should increase, what would happen to the velocity of propagation? Why?

27-9. Non-Resonant Line

A NON-RESONANT LINE is defined as one of infinite length or one that is terminated with a resistive load equal in ohmic value to the characteristic impedance of the line. In a non-resonant line, all of the energy transferred down the line is absorbed by the load resistance and any inherent resistance in the line. The voltage and current waves are called TRAVELING WAVES, and move in phase with one another from the source to the load. On lines carrying radio frequency energy, the line is almost always terminated with a resistance or impedance equal to the characteristic impedance of the line.

Since the non-resonant line may be either an infinite line or one terminated in its characteristic impedance, the physical length of the line is not critical. In the resonant line that will be discussed shortly, the physical length of the line is quite important.

The circuit in Figure 27-14 shows a line terminated with a resistance equal to its characteristic impedance. The charging process, and the ultimate development of a voltage across the load resistance will now be described.

At the instant switch S_1 is closed, the total applied voltage is felt across inductor L_1. After a very short time has elapsed, capacitor C_1 begins to assume a charge. C_2 cannot charge at this time because all of the voltage felt between point c and d is developed across inductor L_2 in

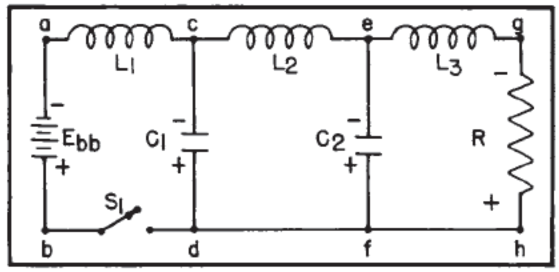


Figure 27-14 - Charged non-resonant line.

the same way as the initial voltage was developed across inductor L_1. Capacitor C_2 is unable to charge until the charge on C_1 approaches the amplitude of the supply voltage. When this happens, the voltage charge on capacitor C_2 begins to rise. The voltage across capacitor C_2 will be felt between points e and f. Since the load resistor is also effectively connected between points e and f, the voltage across the resistor will be equal to the voltage appearing across C_2. The voltage input has been transferred from the input to the load resistor. While the capacitors were charging, the ammeter recorded a current flow. After all of the capacitors are charged, the ammeter will continue to indicate the load current that will be flowing through the dc resistance of the inductors and the load resistor. The current will continue to flow as long as switch S_1 is closed. When it is opened, the capacitors will discharge through the load resistor in much the same way as filter capacitors discharge through a bleeder resistor.

There is little difference between the charging of the line when an ac voltage is applied to it. The charging sequence of the line, with an ac voltage applied, will now be discussed. Refer to the circuit and waveform diagrams in Figure 27-15.

As the applied voltage begins to go positive, the voltage wave begins traveling down the line. At time T_3, the first small change in voltage arrives at point a, and the voltage at that point starts increasing in the positive direction. At time T_5, the same voltage rise arrives at point b, and at that time (T_7), the same voltage rise arrives at the end of the line. The waveform has moved down the line and does so as a wavefront. The time required for the voltage changes to move down the line is the same as the time required for the dc voltage to move down the same line. The time for both of these waves to move down the line for a specified length may be computed by using equation 27-4. The following general remarks concerning the ac charging of the line may now be made: all of the instantaneous voltages produced by the generator travel down the line in the order in which

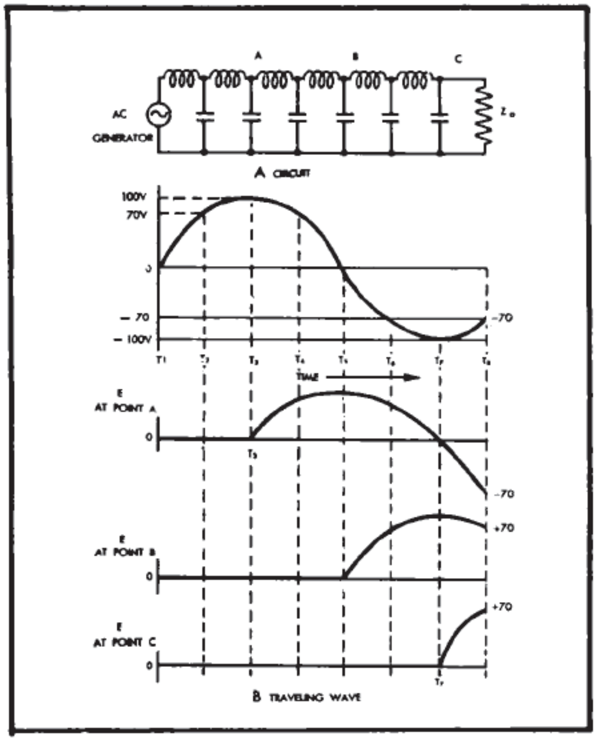


Figure 27-15 - AC charging analysis.

they were produced. If the voltage waveform is plotted at any point along the line, the resulting waveform will be a duplicate of the generator waveform. Since the line is terminated with its characteristic impedance, all of the energy produced by the source will be absorbed by the load impedance.

Q3. What is the relationship between the velocity of propagation for a given line segment with equal values of ac and dc voltages applied? Explain.

27-10. Resonant Transmission Line

A RESONANT LINE is defined as a transmission line that is terminated with an impedance that is NOT equal to its characteristic impedance. Unlike the non-resonant line, the length of the resonant line is critical. In some applications, the resonant line may be terminated in either an open or a short. When this occurs, some very interesting effects may be observed.

A transmission line of finite length terminated in an open circuit is illustrated in Figure 27-16. The characteristic impedance of the line may be assumed to be equal to the internal impedance of the source. This is done to assure maximum power transfer. Even though the impedance of an open circuit is often assumed to be infinite, as it should be, a characteristic impedance for this type of line is specified. The only difference between this line and the infinite line previously discussed is that this line has a finite length.

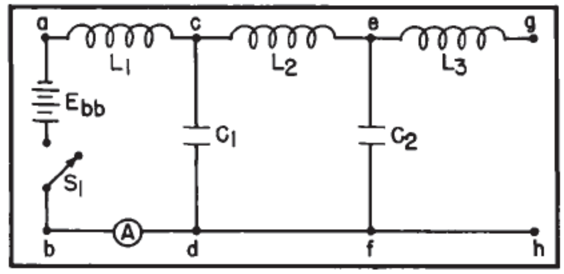


Figure 27-16 - Open-ended transmission line.

Since the impedance of the source is equal to the impedance of the line, the applied voltage will be divided equally between the impedance of the source and the impedance of the line. When switch S_1 is closed, current begins to flow as the capacitors begin to charge through the inductors. As each capacitor charges in turn, the voltage will move down the line. As the last capacitor is charged to the same voltage as every other capacitor, there will be no difference in potential between points e and g. This is true because the capacitors will possess exactly the same charge. The inductor L_3 is also connected between points e and g. Since there is no difference in potential between points e and g, there can no longer be current flow through the inductor. This means that the magnetic field about the inductor will no longer be sustained. The magnetic field must collapse. It is characteristic of the field about an inductor to tend to keep current flowing in the same direction when the magnetic field collapses. This additional current must flow into the capacitive circuit of C_3. Since the energy stored in the magnetic field is equal to that stored in the capacitor, the charge on capacitor C_3 will double. The voltage on capacitor C_3 will be equal to the value of the applied voltage. Since there is no difference of potential between points c and e, the magnetic field about inductor L_2 will collapse, forcing the charge on capacitor C_2 to double its value. The field about inductor L_1 will also collapse, doubling the voltage on C_1. The combined effect of the collapsing magnetic field about each inductor in turn causes a voltage twice the value of the original to apparently move back toward the source. This voltage movement in the opposite direction caused by the conditions just described is called REFLECTION. The reflection of voltage occurred in the same polarity as the original charge. Therefore, it is said of a transmission line terminated in an open, that the reflected

A2. The velocity of propagation will decrease because the time required to propagate the front will increase.

A3. The time required is the same. The time is dependent on the values of L and C and not on the magnitude or type of voltage applied.

wave will always be of the same polarity and amplitude as the original voltage wave. When this reflected voltage reaches the source, the action stops because of the cancellation of the voltages. The current, however, is reflected back with an opposite polarity because when the field about the inductor collapsed, the current dropped to zero. As each capacitor is charged, causing the reflection, the current flow in the inductor that caused the additional charge drops to zero. When capacitor C_1 is charged, current flow in the circuit stops and the line is charged. It may also be said that the line now "sees" that the impedance at the receiving end is an open.

When a sine wave of voltage is applied to the line, the sine wave initially transmitted down the line is called an INCIDENT WAVE. A reflection will also occur when a sine wave is applied to the line. However, this reflection is more appropriately called a REFLECTED WAVE. If a sine wave of 50 volts is applied to the line, the reflected wave will be in phase at a voltage magnitude of 25 volts. The sum of the incident (25 V) plus the reflected wave (25 V) equals 50 V or the applied voltage. The initial current wave, whatever its value, will be reflected back out of phase, but equal in magnitude to the original current wave. The effects realized when the load end is shorted will be analyzed using the diagram in Figure 27-17.

Figure 27-17 illustrates a transmission line that is terminated in a short. As the switch is closed, the source sees the characteristic impedance of the line. Let it be assumed that the characteristic impedance of the line is equal to the internal impedance of the source for the

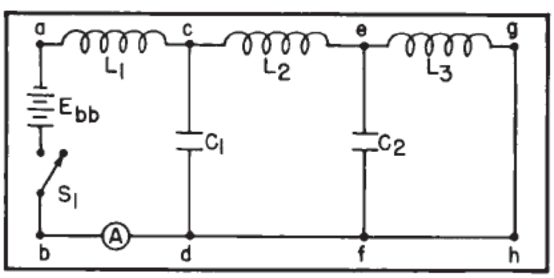


Figure 27-17 - Shorted transmission line.

purpose of maximum power transfer. When the switch is closed, capacitor C_1 begins to charge through inductor L_1 in much the same way as in the infinite line. As the charge on C_1 increases to a value near the value of the applied voltage, capacitor C_2 begins to charge through inductor L_2. As the wavefront reaches points e and f, there is, due to the short between points g and h, a radical change in the impedance. The impedance to the right of points e and f will be much less than the characteristic impedance of the line. The current in that line segment will increase. The electron supply for this increased current will come from capacitor C_2. C_2 will discharge, making point e less negative than it had been. C_2 will discharge to zero volts. This discharge current will flow through inductor L_3. Since the energy contained in the capacitor is equal to the energy contained in the inductor, the current flow in the inductor will double. Since the current flow through inductor L_3 must be the same as the current flow through inductor L_2 (they are in series), capacitor C_1 will discharge through L_2 increasing the current flow through it. Point c with respect to point d will become more positive as capacitor C_1 discharges. Capacitor C_1 will discharge down to zero volts thus reflecting back to the source the zero volt condition caused by the short between point g and h. This voltage reflected back to the source is of an opposite polarity to the original voltage. This can be easily shown by the fact that as the original voltage was transmitted down the line, point c became more negative with respect to point d as capacitor C_1 assumed a charge. However, upon reaching the short, the capacitors were caused to discharge causing point c to go positive with respect to point d. It may then be said that the voltage reflected from a line terminated in a short is of equal amplitude but opposite in polarity.

The current flowing in the circuit may be either said to be of the same polarity, or if ac is used, in phase. The current will have doubled conforming to the equation $2V/Z_o$. When the reflected wave travels back to the source, it being of opposite polarity to the original will cause a zero condition to exist at the battery terminals because of cancellation. The dc source will then send out a new wave of voltage with a current equal to V/Z_o making the total current now equal to $3V/Z_o$. On the reflection of the second wave of voltage, the current builds to a value of $5V/Z_o$. This increasing current will continue to build in this manner, until the current handling capacity of the line is overcome. This high value of current seems to be the logical result because if the dc current path of the circuit is examined, the only opposition to the flow of direct current is the internal

resistance of the source, the inherent resistance of the wires, and the low value of dc resistance offered to the circuit by the inductor windings. If an ac voltage is applied to the same line terminated in a short, the effects would be exactly the same.

The conditions of an open or shorted transmission line are extraordinary conditions. In many applications, the transmission line is not only chosen for maximum power transfer, but also for impedance matching. Reflected waves are, on the whole, highly undesirable. It is known that when the impedance of the source is equal to the characteristic impedance of the line, and that line is terminated in its characteristic impedance; there is a maximum transfer of power, complete absorption of energy by the load, and no reflected waves. If the line is not terminated in its characteristic impedance, there will be reflected waves present on the line. The type and amount of reflected waves is dependent upon the type and amount of mismatch. When a mismatch occurs, there is an interaction between the incident and reflected waves. This interaction results in the creation of a new kind of wave called the STANDING WAVE. The name standing wave is given to these waveforms because they apparently remain in one position, varying only in amplitude. These waves, and the variations in amplitude, are illustrated in Figure 27-18.

All of the wave diagrams in Figure 27-18 show the relationship that exists between the incident and the reflected wave in a transmission line that is terminated in an open. The dark line that is shown superimposed on the waveforms represents the standing waves. Standing waves are the instantaneous sum of both the incident and the reflected waves. In the open transmission line, the relationship between the incident and the reflected wave is such that they are equal in amplitude and phase at the receiving end. The phase relationships in Figure 27-18 may appear confusing, but it must be remembered that the incident wave is moving to the right while the reflected wave is moving to the left. In each of the diagrams in Figure 27-18, the instantaneous sum (standing waves) is plotted using the heavy dark line. In diagrams two and six, the waveforms coincide, and at that point and time, the voltage is zero. If the diagrams are examined further, it is found that at a point one quarter of the distance from the end, and at a point three quarters the distance from the same end, the voltage is zero at all times. Because of the zero stationary point, the waves are appropriately called standing waves.

If an ac meter was used, and current measurements were taken on the line, the current readings would only be of magnitude and not polarity. If the values thus found were plotted on a graph, the current curve would appear as the positive going waves illustrated in Figure 27-19. This is the conventional picture of standing waves.

At the end of a transmission line terminated in an open, the current is zero and the voltage is maximum at the terminating end. This relationship may be stated in terms of phase. The voltage and current at the end of an open ended transmission line are 90° out of phase.

At the end of a transmission line terminated in a short, the current is maximum and the voltage is zero. The voltage and current are again 90° out of phase.

These current-voltage relationships are shown in the diagrams in Figure 27-20. These phase relationships are important because they will indicate how the line will act at different points throughout its length.

A transmission line will have points of maximum and minimum voltage as well as points of maximum and minimum current. The position of these points can be accurately predicted if the applied frequency and type of line termination are known.

A wave that is radiated through space travels at a speed of approximately 186,000 miles per second (300 million meters per second). The velocity of this wave is constant regardless of frequency, so that the distance traveled by the wave during a period of one cycle (called one wavelength) can be found by the formula:

$$\lambda = \frac{V}{F} \qquad (27\text{-}6)$$

$$\lambda = \frac{300,000,000}{f}$$

where λ (the Greek letter lambda, used to symbolize wavelength) is the distance in meters from the crest of one wave to the crest of the next, f is the frequency in cycles per second, and 300,000,000 is the velocity of the radio wave in meters per second. It should be noted, however, that the wavefront travels more slowly on a wire than it does in free space. Conversion factors must be used when calculating the distance energy travels in a transmission line during one cycle, or wavelength. The conversion factor, or constant, for a parallel line is 0.975 and for an air insulated concentric (coaxial) line is 0.85.

A resonant line, like any tuned circuit, will be resonant at a particular frequency. The line will present to its source, a high or low resistive impedance at multiples of a quarter wave length depending on termination at output end.

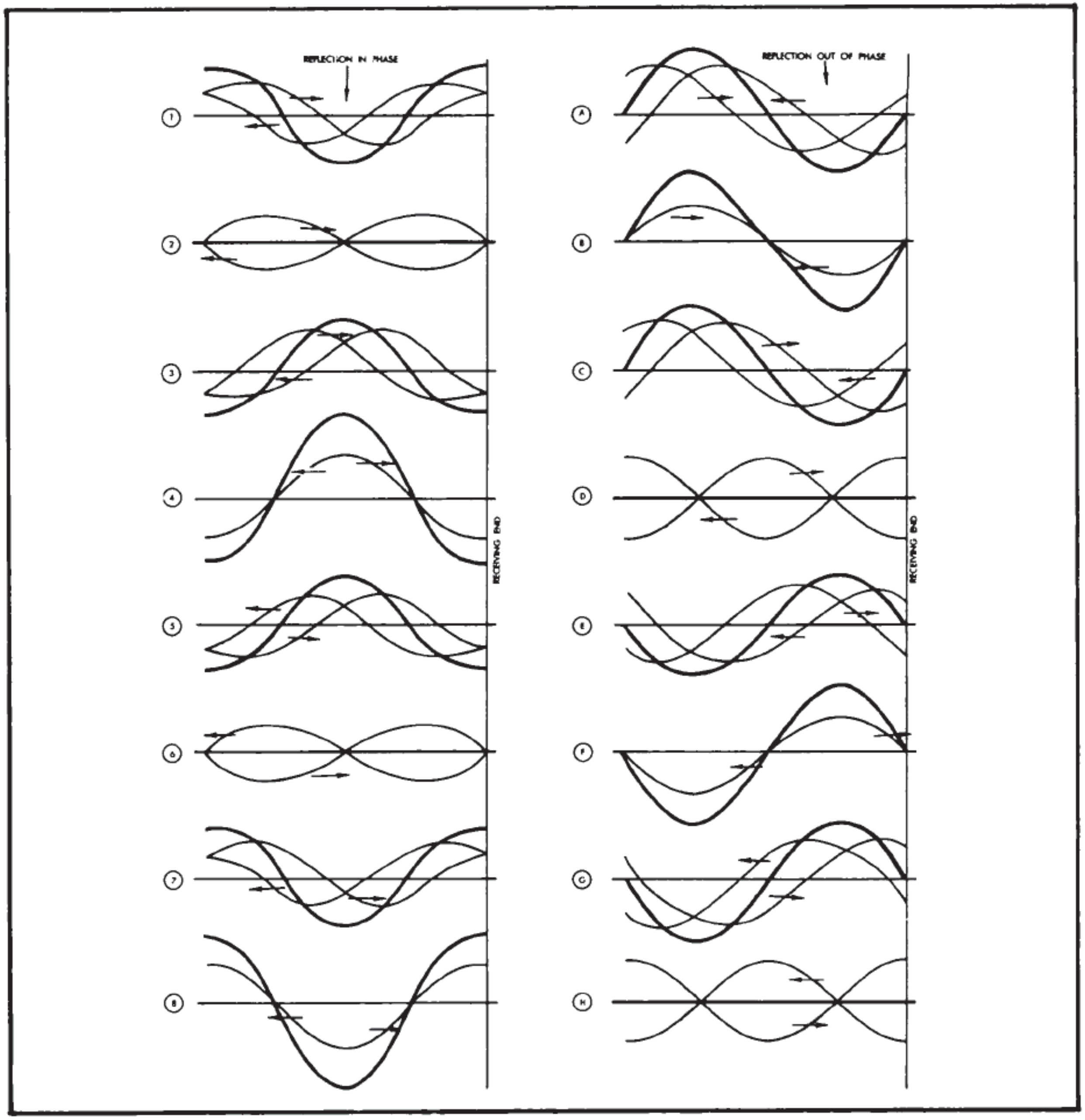


Figure 27-18 - Standing waves.

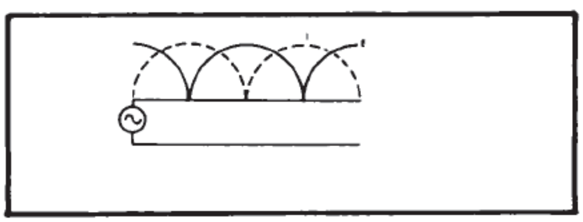


Figure 27-19 - Conventional picture of standing waves.

At points that are not exact multiples of a quarter wave length, the line acts as a capacitor or an inductor. The reasons for this action will be pointed out shortly.

A resonant transmission line may assume the characteristics of a resonant circuit composed of lumped inductance and capacitance. The more important circuit effects that resonant transmission lines have in common with resonant circuits having lumped inductance and capacitance are as follows:

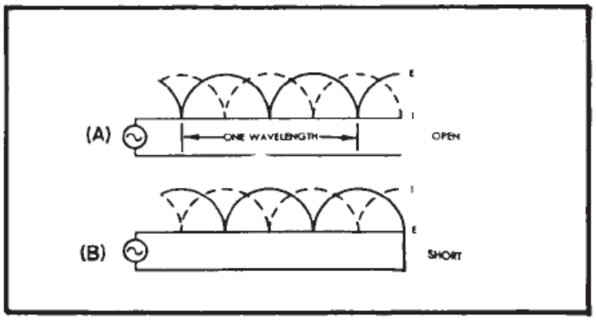


Figure 27-20 - Standing waves of voltage and current. (a) line terminated in an open. (b) line terminated in a short.

Series resonance - resonant rise of voltage across the reactive circuit elements, and a low impedance across the resonant circuit.

Parallel resonance - resonant rise of current in the reactive circuit elements, and a high impedance across the resonant circuit.

The open ended resonant line may be better understood by a consideration of the current, voltage, and impedance curves as shown in Figure 27-21.

The transmission line in Figure 27-21 and the transmission lines to follow will be assumed to be loss free unless specified otherwise. If losses are present, the voltage at the node point is not zero, and the current at the current node point is not zero. Figure 27-21 illustrates the relationship between voltage, current, and impedance for various lengths of open-ended transmission lines. The impedance that the generator "sees" at various distances from the output end is shown directly above the impedance curves. The curves above the letters (R, X_L, X_C) of various heights indicate the relative magnitudes of the impedances presented to the generator for the various lengths of line indicated. The letters themselves indicate the type of impedance offered at the corresponding inputs. The circuit symbols above the various transmission lines indicate the equivalent electrical circuits for the transmission line at that particular length (measured from the output end). The curves of effective E and I whose ratio (E/I) is the impedance Z, are shown above each line.

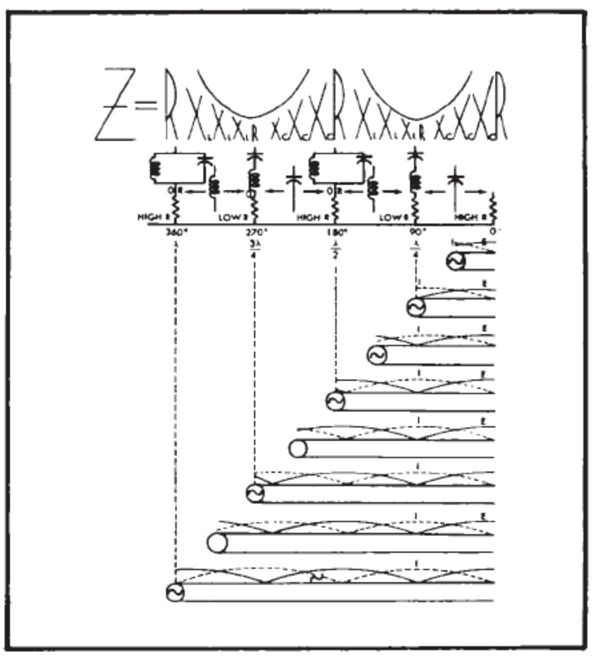


Figure 27-21 - Impedance characteristics of an open-ended transmission line.

At all ODD quarter wavelength points ($\lambda/4$, 3 $\lambda/4$, 5 $\lambda/4$, etc.) measured from the open end of the line, the current is maximum and the impedance is minimum. In addition, there is a resonant rise of voltage from the odd quarter wavelength points toward the open end. Thus, at all odd quarter wavelength points the open ended transmission line acts like a series resonant circuit. The impedance is therefore very low and is prevented from being zero only by the small losses that may be present.

At all EVEN quarter wavelength points ($\lambda/2$, λ, 3 $\lambda/2$, etc.) the voltage is maximum, and therefore, the impedance is maximum. A comparison of this type of transmission line with an LC resonant circuit shows that at even quarter wavelengths (from the output end) the line acts like a parallel resonant circuit.

In addition to acting as series or parallel LC resonant circuits, resonant open ended lines also may act as nearly pure capacitances or inductances when the lengths of the lines are not exact multiples of the fundamental quarter wavelength corresponding to the frequency of the applied voltage at the input terminals. Figure 27-21 shows that an open end line less than a quarter wavelength long acts like a capacitor between one quarter and half a wavelength, as an inductor between one half and three quarters wavelength, as a capacitance between three quarters and one wavelength, and so forth. These characteristics are set forth by observing the relationships between current and voltage at different points on the line.

The closed end transmission line may also be considered by use of the diagram shown in Figure 27-22. At ODD quarter wavelengths from the closed end of the line, the voltage is high, the impedance is high, and the current is low. Because conditions are similar to those in a parallel resonant circuit, the shorted transmission line of odd quarter wavelengths acts

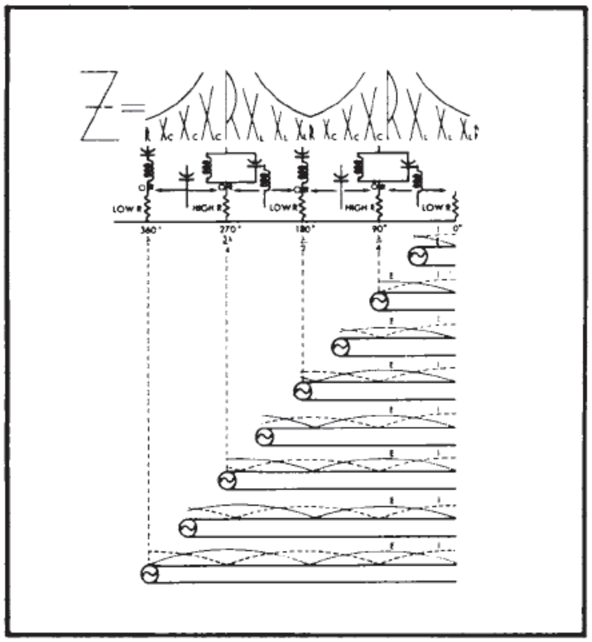


Figure 27-22 - Impedance characteristic of a closed-end resonant line.

like a parallel resonant circuit. The voltage across the circuit cannot exceed the applied voltage.

At EVEN quarter wavelength points (measured from the shorted end), the voltage is minimum, the current is maximum, and the impedance is minimum. Because this action is similar to the action of a series resonant LC circuit, a shorted transmission line of even quarter wavelengths acts like a series resonant circuit.

Resonant closed lines, like open ended lines, may also act like nearly pure capacitances and inductances when the length of the line is not an exact multiple of the fundamental quarter wavelength corresponding to the frequency of the applied voltage at the input terminals.

By choosing a transmission line of particular length, it can be made to perform the function of an inductance, capacitance, or resistance as they may be found in any circuit configuration.

If the frequency applied to the transmission line would vary, the characteristics of the line would change. If the frequency increases, the wavelength becomes shorter. This means that the length of the line at this new frequency is greater than one wavelength. If the wavelength increases to a value of 1 1/4 wavelength, the change in its characteristics would be the same as for a 1/4 wavelength. Of course, the termination of the line is still important.

Q4. What is the relationship between the incident wave and the reflected wave in an infinitely long transmission line?

Q5. What is the relationship between current and voltage at the three quarter wavelength point of a line one wavelength long that is terminated in a short?

Q6. What are the characteristics, at the input terminals, of a transmission line one half wavelength long that is terminated in an open?

27-11. Transmission Lines as Impedance Matching Devices

Because a transmission line has an impedance that varies over its entire length, regardless of the type of termination used, it is well suited for use as an impedance matching device. The impedance of a 1/4 wave section of transmission line shorted at one end varies widely over its length, as indicated by the diagram shown in Figure 27-23. At the shorted end, the current is high and the voltage is low. The impedance of the shorted end is low. At the open end, the conditions are reversed. The current is low, the voltage is high, and the impedance is high. When this 1/4 wave section is used, it is possible to match almost any impedance somewhere along the line. For example, a 300 ohm line may be matched to a 70 ohm line without the creation of standing waves on either of the two lines to be matched. Figure 27-23B shows how the connections are made.

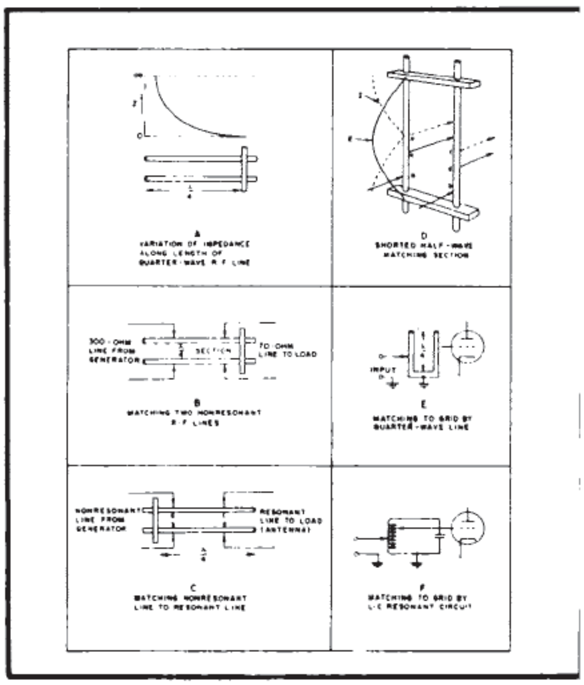


Figure 27-23 - Transmission lines used for impedance matching.

Energy from the 300 ohm line sets up standing waves on the quarter wave section. The connection made between the 300 ohm line and the quarter wave matching section is made at a point where the impedance of the quarter wave section is 300 ohms. The 70 ohm line is similarly adjusted to bring about an impedance match at the shorted end.

The quarter wave line may also be used to match a non-resonant line to a resonant line. This is illustrated in Figure 27-23C. In order to be classified as a non-resonant line, a line must be terminated in its characteristic impedance, and the terminating impedance should be a pure resistance. The impedance of a shorted quarter wave section at the shorting bar is zero, and the impedance increases along the line toward the open end. The shorting bar should be adjusted to make a maximum voltage appear between points c and d, and the contacts, ab, between the non-resonant line and the quarter wave section, are adjusted for the best match.

A half wave section of line shorted at both ends is also used as an impedance matching device, particularly in antenna coupling problems. Figure 27-23D shows a half wave section excited at ab, and having resonant current and voltage values as shown by the curves labeled E and I. The input (from the generator) to ab "sees" an impedance Z_{ab}, equal to the E/I ratio at that point, and the output (load) looking into cd "sees" a larger E/I ratio, hence a larger impedance, Z_{cd}. The greatest impedance will be obtained at ef (where the voltage is highest and the current is lowest). Conversely, the lowest impedance will be at the shorting bars where the current is high and the voltage is low. Because the upper half of the wave section, or half wave frame, repeats the impedance of the lower half. There will always be two points on the frame that possess the same impedance. There will be a difference in the phase of the two currents involved, the current in one half being 180° out of phase with the current in the other half.

Another example of the shorted quarter wave section used as an impedance matching device is shown in Figure 27-23E. In this figure, a relatively low impedance input is transformed to a high impedance to match the high impedance input of the tube. The equivalent circuit is shown in Figure 27-23F.

A non-shorted transmission line may also be used as an impedance matching device. This is shown in Figure 27-24. A non-shorted quarter wave transmission line having the correct impedance may be used to match two dissimilar impedances. The necessary characteristic impedance, Z_o, of the quarter wave matching section is given by the following equation:

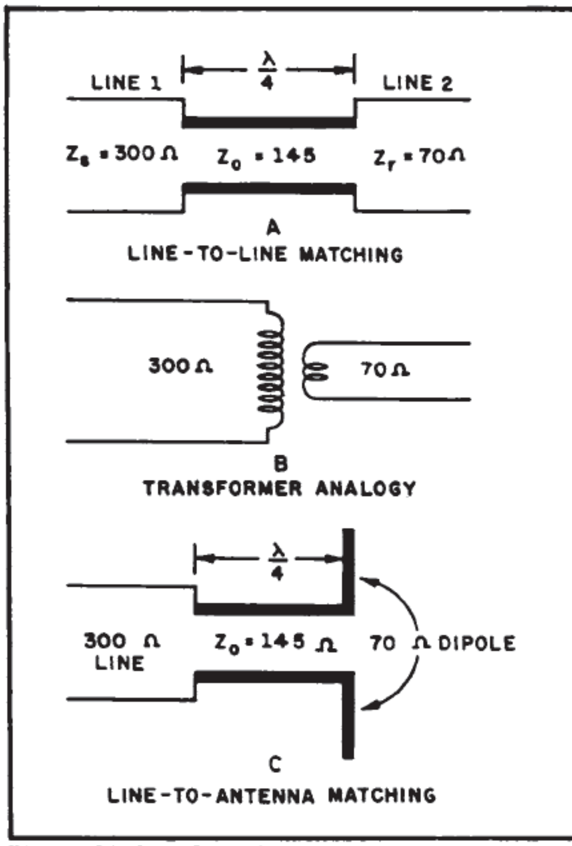


Figure 27-24 - Impedance matching with an unshorted quarter wave line.

$$Z_o = \sqrt{Z_s Z_r} \qquad (27\text{-}7)$$

where: Z_s = the impedance of line one
Z_r = the impedance of line two

A transformer analogy and the means used to connect a 300 ohm line and a quarter wave matching section is shown.

Q7. Why can the transmission line be used as an impedance matching device?

27-12. Transmission Line Sections as Filters

The characteristics of a quarter wave section of a transmission line allow it to become an efficient filter or suppressor of even harmonics. Other types of filters may be used to filter out the odd harmonics. In fact, filters may be designed to eliminate the radiation of an entire single side-band of a modulated carrier.

Suppose that a transmitter is operating on a frequency of 5Mc, and it is found that the transmitter is causing interference on 10Mc and 20Mc. In addition to other uses, the line section may be used to eliminate these undesirable harmonics.

A4. There is no relationship between them because on an infinitely long transmission line, there are no reflected waves.

A5. They are 90° out of phase, the voltage is maximum and the current is minimum.

A6. The impedance is maximum, the current is minimum, and the voltage is maximum. The half wavelength point displays the characteristics of a parallel resonant circuit.

A7. Because the transmission line has an impedance value that varies widely over its entire range.

A quarter wave line shorted at one end offers a high impedance at the unshorted end to the fundamental frequency. At a frequency twice the fundamental, such a line is a half wave line, and at a frequency four times the fundamental, the line becomes a full wave line. A half wave or full wave line offers zero impedance when its output is terminated in a short. Therefore, the radiation of even harmonics from the transmitter antenna can be eliminated almost completely by the circuit shown in Figure 27-25.

The resonant filter line, ab, is a quarter wave in length at 5Mc , and offers almost infinite impedance at this frequency. At the second harmonic, 10Mc , the line ab is a half wave line and offers zero impedance at the antenna, thereby shorting this frequency to ground. At 20Mc the line is a full wave section and it offers little opposition between the antenna and ground. The quarter wave filter may be inserted anywhere along the non-resonant transmission line with the similar effect.

Both open and closed quarter wave resonant lines may be used as filters. This is shown in Figure 27-26.

In this application, a quarter wave filter, b, that is open at the output end is inserted in series with the transmission line. A quarter wave line that is open at the output end offers

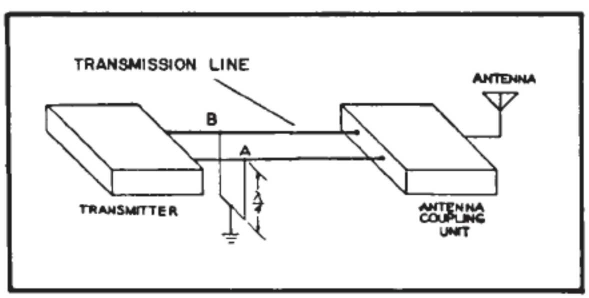


Figure 27-25 - Quarter wave filters.

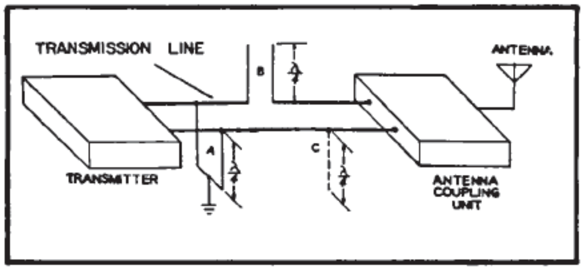


Figure 27-26 - Quarter wave filters.

low impedance at the input end to the fundamental frequency. At each odd harmonic, such a line is an odd multiple of a quarter wave, and therefore, offers little impedance to odd harmonics. Actually, the impedance at b is so low at the fundamental frequency that it may be considered a continuous line, as if a short were placed across the base of the quarter wave line. Thus, the quarter wave open-filter line, b, in the diagram shown in Figure 27-26, passes the fundamental and the odd harmonics along the line to the antenna coupling unit. At even harmonics, however, the length of the open line (at b) becomes a half wave, or some multiple of a half wave, so that line b offers a high impedance to the even harmonics, and blocks their passage to the antenna coupling unit.

Unfortunately, the foregoing methods of inserting wave filters in shunt with a line cannot be used to eliminate odd harmonics, because any attempt to eliminate the odd harmonics also results in serious loss to the fundamental frequency. For example, assume that the line in Figure 27-26 is a quarter wave at the third harmonic (15Mc). This frequency would be eliminated effectively before it reached the antenna coupling unit. However, the fundamental that is to be transmitted would also be greatly attenuated. If a line is quarter wavelength at 15Mc, it is a twelfth-wave in length at 5Mc (wavelength varies inversely with frequency). A line a twelfth-wave in length would act as a capacitor, and offer low impedance at 5Mc. Therefore, although 15Mc radiation would be suppressed considerably, 5Mc would also be attenuated.

27-13. Standing Wave Ratio

It was previously pointed out that the presence of standing waves is disadvantageous. Their presence indicated a mismatch in the input and output of the systems connected by a transmission line. Because the voltage amplitude of the standing waves may be extremely high, the possibility of insulation breakdown is great. Since the voltage and current magnitudes vary at different points along the line, the power

handling capacity of the line is reduced. Transmission lines are not normally used where the power handling requirement is high. For high power applications, waveguides are used. Since the current flow in the standing waves is high, the copper loss of these lines is considerable. Their efficiency is low for the same reason. Because there is a difference in phase between the current and voltage, the line power factor is also low. Radiation losses are also high when standing waves are present.

To minimize these losses, the impedance of the load and the line should be the same. If they are not the same, standing waves will be produced, and the disadvantages will be apparent. To determine the presence or magnitude of the standing waves, two-wire transmission lines called LECHER LINES are used. The amplitude of the current at both the loops and the nodes may be measured by the use of Lecher Lines. The ratio of the amplitude of the loop current and the node current is known as the STANDING WAVE RATIO. This ratio is also the ratio of the characteristic impedance of the line and the impedance of the load. The desirable standing wave ratio is one or unity. That would mean that all of the energy generated by the source and transmitted by the transmission line would be absorbed by the load. Figure 27-27 illustrates two mismatched transmission lines.

Let it be assumed that the characteristic impedance of the lines in Figure 27-27 is a value of 500 ohms. In the transmission line labeled (A), the load impedance is 250 ohms. This ratio will be 2. In the transmission line labeled (B), the load impedance is equal to 100 ohms. The standing wave ratio will be 5. In general, the higher the SWR (standing wave ratio), the greater the mismatch between line and load. A knowledge of the position of the voltage and current nodes and their values will indicate whether the load impedance is less or greater than the characteristic impedance of the line. Lecher Lines will be used to determine this relationship. In general, if the voltage at the load end of the line is maximum, and the current is minimum, the impedance of the load will be greater than the characteristic impedance of the line. The maximum effective voltage and minimum current are indicative of conditions approaching an open.

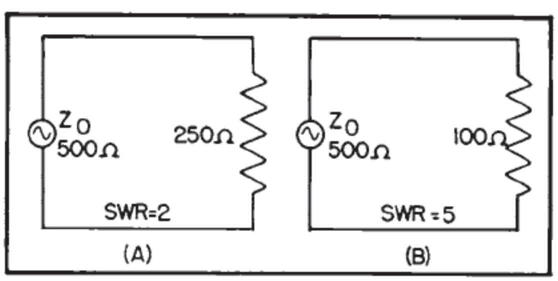


Figure 27-27 - Mismatched lines.

If the current is maximum and the voltage is minimum across the output of the line, conditions are similar to those approaching a short. It is a reasonable assumption that the impedance of the line is greater than the impedance of the load.

LOSSES IN TRANSMISSION LINES

Whenever the electrical characteristics of lines are explained, the lines are often thought of as being loss free. Although this allows for simple and more readily understood explanations, the losses in practical lines cannot be ignored. There are three major losses that occur in transmission lines, they are: copper losses, dielectric losses, and radiation or induction losses.

27-14. Copper Losses

It can be stated that there are two types of copper losses. One type is the I^2R loss. The resistance of any conductor is never zero. When current flows through a transmission line, energy is dissipated in the form of heat. This loss of energy is the I^2R, or power loss. A reduction in resistance will minimize the power loss in the line. The resistance is indirectly proportional to the cross sectional area. Keeping the line as short as possible will decrease the resistance and the I^2R loss. The use of wire with a large cross sectional area is also desirable; however, this method has its limitations since any change in wire size will also cause a change in the characteristic impedance of the line.

Another type of copper loss is due to skin effect. When a dc current flows through a conductor, the movement of electrons through its cross section is uniform. The situation is somewhat different when ac is applied. The expanding and collapsing fields about each electron encircle other electrons. This phenomenon, called self induction, retards the movement of the encircled electrons. Figure 27-28 shows a cross sectional view of a current carrying wire. The flux density at the center is so great that electron movement at this point is reduced. As frequency is increased, the opposition to the flow of current in the center of the wire increases. Current in the center of the wire becomes smaller and most of the electron flow is on the wire surface. When the frequency applied is 100Mc or higher, the electron movement in the center is so small, that the center of the wire would be removed without any notice-

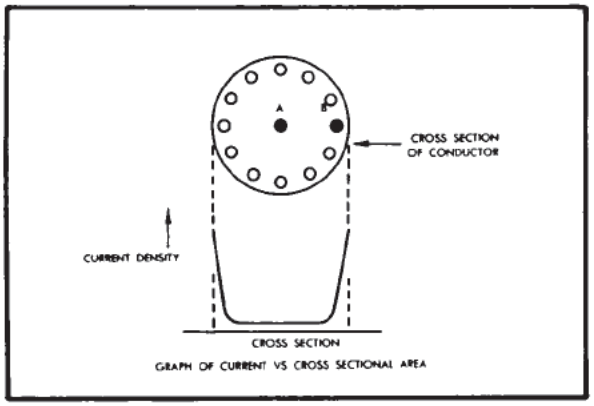


Figure 27-28 - Skin effect losses in RF lines.

able effect on current. It can be seen that the effective cross sectional area decreases as the frequency increases, and since resistance is inversely proportional to cross sectional area, the resistance will increase as the applied frequency increases.

Since the power loss increases as the resistance increases, power losses also increase with an increase in frequency.

The conductivity of the line can be increased by plating it with silver. The plating then becomes the major carrier of current.

27-15. Dielectric Losses

Dielectric losses result from the heating of the insulating material between the conductors. The production of this heat requires power which must be taken from the source. Figure 27-29A shows the normal electron orbital path when there is no difference of potential between the two wires. Figure 27-29B shows how the orbital path of the electrons in the dielectric has been changed. The electron path has been altered because it has been repelled by the negative potential of one wire, and attracted by the positive potential of the other. This distortion of the field requires power from the source and increases the losses. The structure of the atoms of some materials is harder to distort, for example rubber, and more energy is absorbed from the source. The electron path of some atoms is easily altered and requires very little

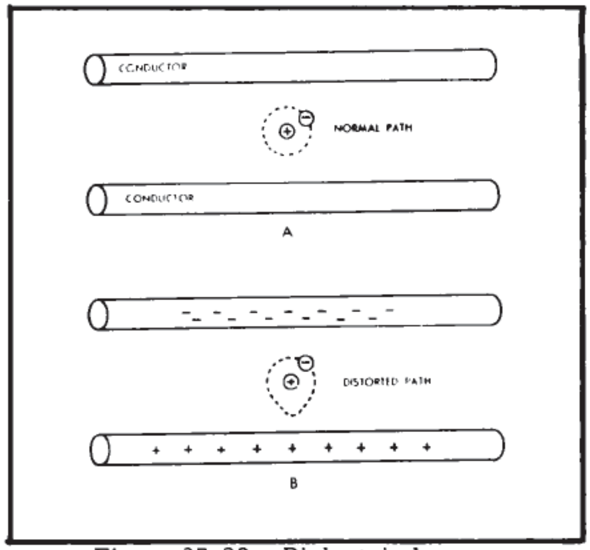


Figure 27-29 - Dielectric losses.

energy from the source. The losses that occur when air is used as the dielectric are very small since little energy is required in the distortion of its atomic structure. In many cases the use of a solid dielectric is required, for example in the flexible coaxial cable, and if losses are to be minimized an insulation with a low dielectric constant is used. Polyethylene allows the construction of a flexible cable whose dielectric losses, though higher than air, are still much lower than the losses that occur with other types of low cost dielectrics.

27-16. Radiation and Induction Losses

The electrostatic and electromagnetic fields that surround a conductor also cause losses in transmission lines. The action of the electrostatic fields is to charge neighboring objects; while the changing magnetic field induces an EMF in nearby conductors. In either case, energy is lost.

Radiation and induction losses may be greatly reduced by terminating the line with a resistive load equal to the lines characteristic impedance, and by proper shielding of the line. Proper shielding can be accomplished by the use of coaxial cables with the outer conductor grounded.

EXERCISE 27

1. Describe the types of transmission lines.
2. Compare the advantages and disadvantages of the parallel two-wire line, the coaxial line, and the waveguide.
3. What is meant by the term "skin effect"?
4. Describe the lumped values found in transmission lines.
5. What does "termination" mean?
6. What does the term "characteristic impedance" mean?
7. What is the impedance of a line which has 6 microfarads of capacitance and an inductive reactance of 269 ohms at the operating frequency of 4.8 kc?
8. If the wires of a two wire line are placed farther apart, what happen to the characteristic impedance of the line?
9. What is meant by the term "propagation"?
10. How does the characteristic resistance of a transmission line compare to the characteristic impedance?
11. What does the term "velocity of propagation" mean?
12. What is a traveling wave?
13. What is the difference between a resonant line and a non-resonant line?
14. What is a reflected wave?
15. What is an incident wave?
16. Describe the characteristics of reflection in both the open and shorted transmission line.
17. Describe the current buildup in the shorted transmission line.
18. What are the phase relationships between the incident and reflected waves in both the open and shorted lines?
19. What is a standing wave?
20. What does SWR mean?
21. What is meant by the terms wavelength, quarter wave, and half wave?
22. What type of impedance may be expected a quarter wavelength from the end of a transmission line that is terminated in a short? In an open?
23. Describe some of the applications of transmission lines.
24. Describe the quarter wave filter.

CHAPTER 28

ANTENNA AND WAVE PROPAGATION

An antenna is a conductor so constructed as to either radiate electromagnetic energy, to collect electromagnetic energy, or to do both. Except those serving both functions, antennas lie in either of two general categories—transmitting antennas and receiving antennas. A transmitting antenna converts electrical energy into electromagnetic waves (called radio waves) which radiate away from the antenna at speeds near the velocity of light. A receiving antenna converts electromagnetic waves which it intercepts into electrical energy and applies this energy to electronic circuits for interpretation.

The electrical and physical features of antennas are determined by the functions they serve. Such features will vary with operating frequency, power handling capability, plane of polarization and desired radiation field pattern. The physical size of an antenna is determined primarily by its operating frequency and power handling capability, while its shape and height are determined by the desired radiation field pattern.

Radiation of electromagnetic energy is based on the principle that a moving electric field creates a magnetic field, and conversely, a moving magnetic field creates an electric field. This creative process is called PROPAGATION and is used to describe the manner in which electromagnetic waves travel through space.

The electric (E) and magnetic (H) fields which comprise electromagnetic radiation are perpendicular to each other and to their direction of motion. The term POLARIZATION, as applied to radio waves, refers to the direction in which the electric field travels with respect to the earth and is primarily a function of antenna orientation with respect to the earth.

When RF currents flow through a transmitting antenna, radio waves are radiated in all directions, spreading out in much the same manner that waves spread out on the surface of a pond into which a stone has been thrown. Antennas which radiate in all directions are referred to as OMNIDIRECTIONAL. Being electromagnetic in nature, radio waves can be directed so that radiation occurs in specific directions or concentrated into a narrow beam. Such DIRECTIONAL antennas form an important part of point-to-point communication systems and radar systems.

Usually, the most important characteristic of an antenna is its directional property or simply its directivity. Directivity means that an antenna radiates more energy in one direction than in another. For that matter, all antennas are directional, some slightly; others, almost entirely. Some antennas are required to send all energy in one direction in order that as much as possible of the electromagnetic energy generated by the transmitter will strike an object in a given direction. In other systems, it is desirable for the energy to be radiated equally well in all directions from the source. An example of an antenna system which radiates energy in a given direction is the airborne navigation and bombing set. In this set, there is only a limited amount of power available at the transmitter. In order to achieve maximum benefit from this minimum power, all of it is sent in the same direction.

Since the antenna in this set is also used for reception, it likewise receives electromagnetic energy only from one direction. Because of design features, it is possible to tell the direction of an object at which this directional type antenna is sending energy, or the direction of the object from which the antenna is receiving energy. Furthermore, the physical position of the antenna is indicative of the direction of the object. An example of a non-directional antenna is the antenna installed in the radar beacon. This antenna must receive energy equally well in all directions in order that a radar equipped airplane can ascertain its position regardless of its direction from the beacon antenna.

In this chapter, information will be subdivided into two major categories—antennas and wave propagation. Under the subject of antennas, the physical and electrical properties of various basic antennas will be considered. Included will be information concerning radiated field patterns and polarization. Under the subject of wave propagation, various components of the radiated wave will be discussed in detail.

ANTENNAS

28-1. Function of an Antenna

As stated previously, most antennas serve either of two functions: the generation or the collection of electromagnetic energy. In a transmitting system, a radio frequency signal is developed, amplified, modulated and applied

774-422 O - 65 - 7

to the antenna. The RF currents flowing through the antenna produce electromagnetic waves which radiate into free space. In a receiving system, electromagnetic waves passing the antenna induce alternating currents for use by the receiver.

To have adequate signal strength at the receiver, either the power transmitted must be extremely high, or the efficiency of the transmitting and receiving antennas must be high because of the high losses in wave travel between the transmitter and the receiver, and the lower sensitivity of the ultra-high frequency range receivers. Since it is difficult to generate large amounts of power efficiently in the higher frequency ranges, it is especially important that as much of the available signal at the transmitter be converted into radiated energy as possible, and that as much of the radiated energy as possible be picked up at the receiver.

Any antenna transfers energy from space to its input terminals with the same efficiency with which it transfers energy from the output terminals into space, assuming that the frequency is the same. This property of interchangeability of the same antenna for transmitting and receiving operations is known as antenna RECIPROCITY—a convenient principle for those systems (such as radar) incorporating both operations.

Antenna reciprocity is possible chiefly because antenna characteristics are essentially the same regardless of whether an antenna is sending or receiving electromagnetic energy. Because of antenna reciprocity, most radar sets use the same antenna both for receiving and for transmitting.

An automatic switch in the radio frequency line first connects the antenna to the transmitter, then to the receiver, depending upon the sequence of operation. Because of reciprocity of radar antennas, this chapter treats antennas from the viewpoint of the transmitting antenna, with the understanding that the same principles apply equally well when the antennas are used for receiving electromagnetic energy.

Not only do antennas produce or collect electromagnetic energy, but they should do so in an efficient manner. Consequently, antennas are composed of conductors arranged in such a fashion as to permit efficient operation.

28-2. The Dipole Antenna (Hertz)

Any antenna having a physical length that is one-half wavelength of the applied frequency is called a HERTZ antenna. Hertz antennas are predominately used with frequencies above 2 mc. It is unlikely that a Hertz antenna will be found in applications below 2 mc because, at these low frequencies, this antenna is physically too large for most applications. Usually at frequencies below 2 mc, a MARCONI type of antenna is used. This is a quarter-wavelength antenna with ground acting as the other quarter wavelength. Consequently, the difference between the two antennas is that the Hertz type does not require a conducting path to ground while the Marconi type does.

When the open two-wire transmission line was discussed, it was found that one of its disadvantages was excessive radiation at high frequencies. Radiation from a transmission line is undesirable since the perfect transmission line would be one which possessed no losses. Although the two-wire transmission line was considered to be an inadequate transmission line at high frequencies, it can become an effective antenna. For this reason, an analysis of the open-ended, quarter-wave transmission line will furnish an excellent background for understanding basic antenna operations. The open-ended quarter-wave transmission line segment is shown in Figure 28-1.

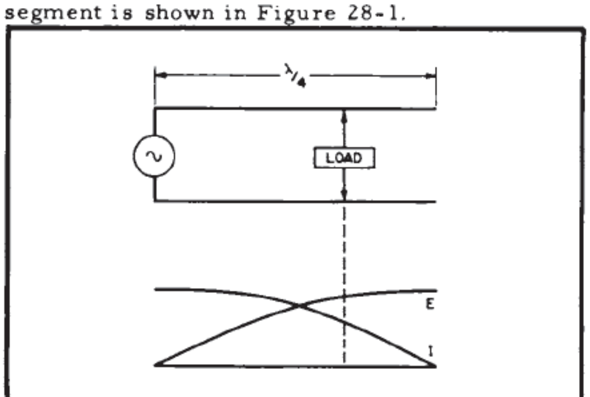


Figure 28-1 - Quarter-wave transmission line segment (open-ended).

The characteristics of the open-ended line are such that the voltage at the end of the line is maximum and the current at the end is zero. This is true of an open-ended line regardless of the wavelength of the line. On either the open or shorted line, standing waves will be produced. Since the voltage applied to the line is sinusoidal, the line will constantly be charging and discharging. Current will be flowing in the line continuously. Since the current at the ends of the line is minimum, a quarter-wave back (at the source) the current must be maximum. The impedance at the sending end is low and the impedance at the receiving end is high. The standing waves of current and voltage are shown on the quarter-wave section in Figure 28-1. Figure 28-2 shows the same quarter-wave line with the impedence curve shown. Also shown in this diagram is the impedance curve for the line.

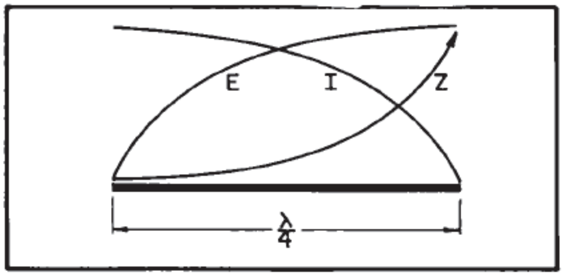


Figure 28-2 - Impedance curve for a quarter-wave line.

Whenever a current is passed through a line, there are fields associated with the flow. There is a magnetic field, and an electrostatic field. The direction of the lines of force which form the magnetic field is determined by using the rule for coils. If the conductor is grasped in the left hand and the thumb is pointed in the direction of current, the fingers surrounding the conductor will point in the direction of the lines of force. The direction of the magnetic lines of force is shown in Figure 28-3A.

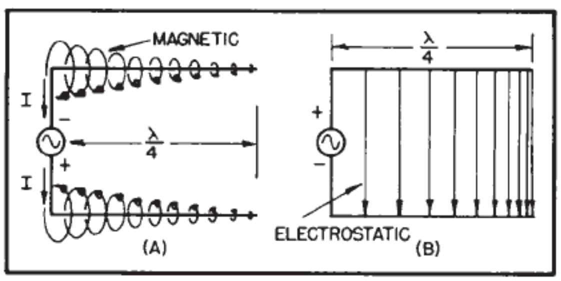


Figure 28-3 - Lines of force about a quarter-wave section.

Notice that the magnetic fields for the two conductors are opposing. This is true because the current is opposite through the two conductors.

An electrostatic field will also be formed between the quarter-wave lines. It will be formed in a direction perpendicular to the charged body with which it is associated. By convention, the direction of the lines of force on a charged body is such that the arrows between the lines point away from the positively charged body and toward the negatively charged body. These lines are shown in Figure 28-3B.

In Figure 28-3, the relative intensities of both fields are shown. The intensity of the magnetic field is strongest where the current is the highest; that is, near the source. Since current decreases as it moves toward the end of the line, the intensity of the magnetic field also decreases. At the end of the line, the magnetic field is minimum because current is minimum.

The voltage is greatest between the open ends of the quarter-wave two-wire line. Therefore, the electrostatic field will be greatest at the open end and minimum at the source end. Although some radiation occurs from the two-wire transmission line at high frequencies, the line cannot provide the efficient radiation expected from a well designed antenna.

It is desirable to have maximum radiation from an antenna. Under such conditions all energy applied to the antenna would be converted to electromagnetic waves and radiated. This maximum radiation is not possible with the two-wire line because the magnetic field surrounding each conductor of the line is in a direction that opposes the lines of force about the other conductor. Under these conditions, the quarter-wave transmission line proves to be an unsatisfactory antenna; however, with only a slight physical modification, this section of transmission line can be transformed into a relatively efficient antenna. This transformation is accomplished by bending each line outward 90° to form a HALF-WAVE or HERTZ ANTENNA, as shown in Figure 28-4.

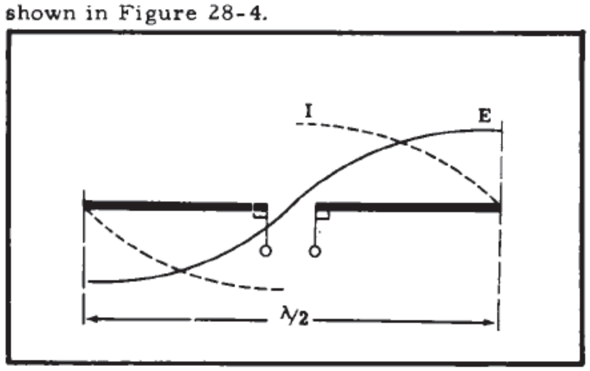


Figure 28-4 - The basic Hertz antenna.

The antenna shown in Figure 28-4 is composed of two quarter-wave sections. The electrical distance from the end of one line to the end of the other is a half wavelength. If a voltage is applied to the line causing current to flow, the current will still be maximum at the source end and minimum at the open end. The voltage will be maximum between the open ends and minimum between the source ends.

An impedance value may be specified for a half-wave antenna thus constructed. Generally, the impedance at the open ends is maximum while that at the source ends is minimum. Consequently, the impedance value varies from a minimum value at the generator to a maximum value at the open ends. An impedance curve for the half-wave antenna is shown in Figure 28-5. Notice that the line has different impedance values for different points along its length. Typical impedance values for half-wave antennas vary from 2500 ohms at the open ends to 73 ohms at the source ends.

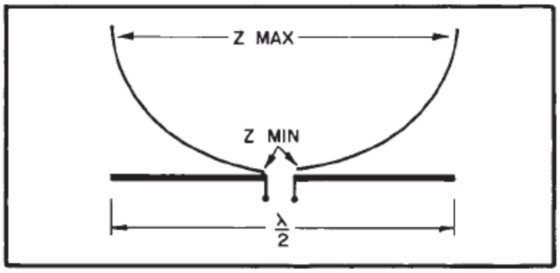


Figure 28-5 - Impedance curve for a half-wave antenna.

Q1. What is the phase relationship between current and voltage along a quarter-wave transmission line?

Q2. What determines the value of impedance along a quarter-wave transmission line?

28-3. Antenna Feedlines

If energy is applied at the geometrical center of an antenna, the antenna is said to be CENTER-FED. If energy is applied to the end of an antenna, it is known as an END-FED antenna. Although energy may be fed to an antenna in various ways, most antennas are either VOLTAGE-FED or CURRENT-FED. When energy is applied to the antenna at a point of high circulating current, the antenna is current-fed. When the generator energy is applied to a point on the antenna of high voltage, the antenna is voltage-fed. Both of these types of feed are shown in Figure 28-6.

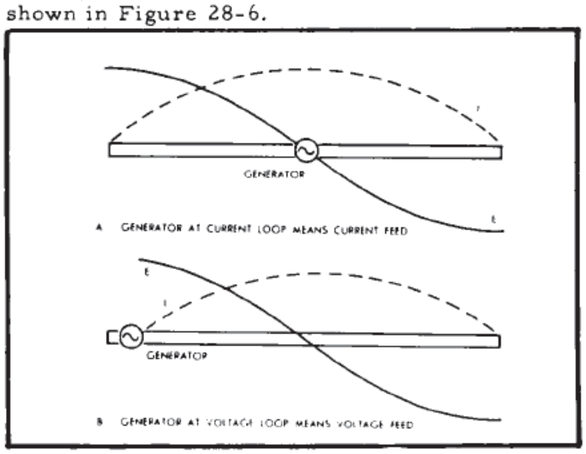


Figure 28-6 - Current and voltage feed.

It is seldom possible to connect a generator directly to an antenna. It is usually necessary to transfer energy from the generator to the antenna by use of a transmission line (also called an antenna FEEDLINE). Such lines may be resonant, nonresonant, or a combination of both types.

The resonant transmission line is not widely used as an antenna-feed method because it tends to be inefficient and is very critical with respect to its length for a particular operating frequency. However, in certain high-frequency applications, resonant feeders sometimes prove convenient.

In a voltage-fed, half-wave antenna energized by a resonant transmission line (as shown in Figure 28-7) one end of the antenna is connected to one side of the transmission line. Voltage changes at that point on the line excite the antenna into oscillation. The impedance at the end of the antenna is very high, since at that point the voltage is high and the current is low ($Z = E/I$).

The transmission line in Figure 28-7 is a half-wavelength, two-wire line. The impedance of such a line is low in order to insure that the high impedance in which it is terminated will produce standing waves. In addition, it is always a multiple of quarter-wavelengths, electrically, so that the line will act as a resonant circuit. This makes the input to the line a high impedance. The parallel resonant circuit composed of the secondary of the transformer and the variable capacitor will develop a high voltage that will then be applied across the high input impedance of the line. Small irregularities in the transmission line length can be compensated for by tuning the variable capacitor.

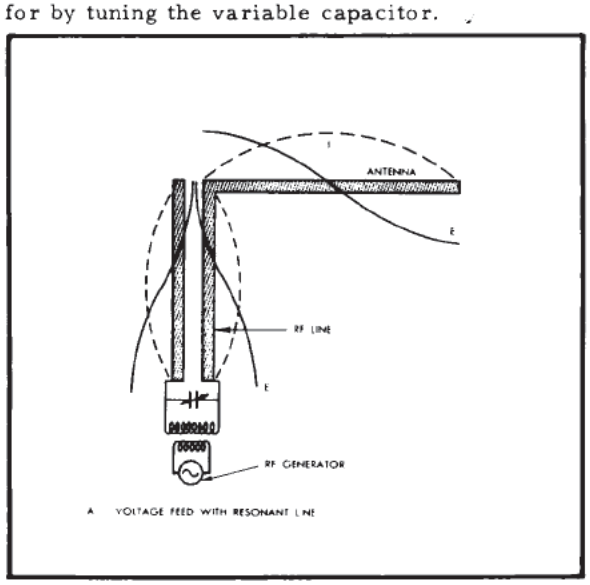


Figure 28-7 - Voltage feed with resonant line.

In the current-fed antenna with a resonant line, as shown in Figure 28-8, the transmission line is connected to the center of the antenna. This antenna has a low impedance at the center and, like the voltage-feeding transmission line, has standing waves on it. Constructing it to be exactly a half wavelength causes the impedance

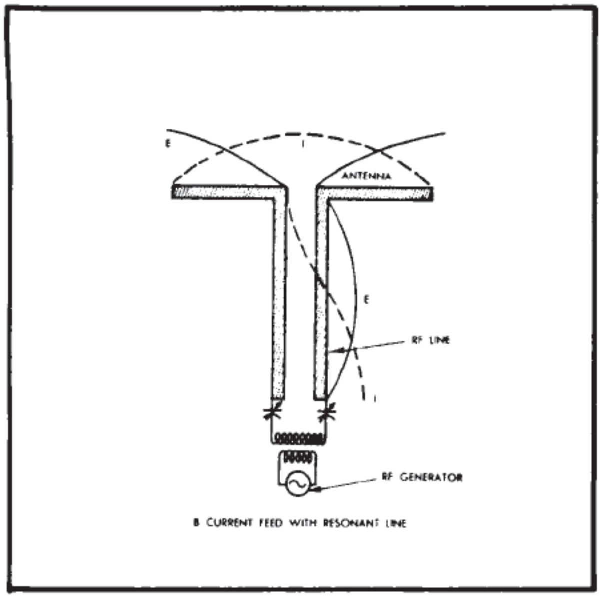


Figure 28-8 - Current feed with resonant line.

at the sending end to be low. A series resonant circuit is used to develop the high currents needed to excite the line. Adjusting the capacitors at the input compensates for irregularities in line and antenna length.

Although these examples of antenna feed systems are simple ones, the principles described apply to antennas and to lines of any length provided both are resonant. The line connected to the antenna may be either a two-wire or coaxial line. In high frequency applications, the coaxial line is preferred.

One advantage of connecting a resonant transmission line to an antenna is that it makes impedance matching unnecessary. In addition, it makes it possible to compensate for any irregularities in either the line or the antenna by providing the appropriate resonant circuit at the input. Its disadvantages are: increased power losses due to high standing waves of current, increased probability of arc-over because of high standing waves of voltage, very critical length, and production of radiation fields by the two-wire line due to the standing waves on it. Radiation fields in a transmission line are undesirable because they represent a loss.

The nonresonant feedline is the more widely used. The open-wire line, the shielded pair, the coaxial line, and the twisted pair may be used as nonresonant lines. This type of line has negligible standing waves if it is properly terminated in its characteristic impedance at the antenna end. It has a great advantage over the resonant line in that its operation is practically independent of its length.

The illustrations in Figure 28-9 show the excitation of a half-wave antenna by nonresonant lines. If the input to the center of the antenna in (A) is 73 ohms, and if a coaxial line with the highest Q has a characteristic impedance of 73 ohms, a common method of feeding this antenna would be through a coaxial cable connected directly to the center of the antenna. This method of connection produces no standing waves on the line when the line is matched to a generator. Coupling to a generator is usually made through a simple untuned transformer secondary.

Another method of transferring energy to the antenna is through the use of a twisted-pair line, as shown in part B of Figure 28-9. It is

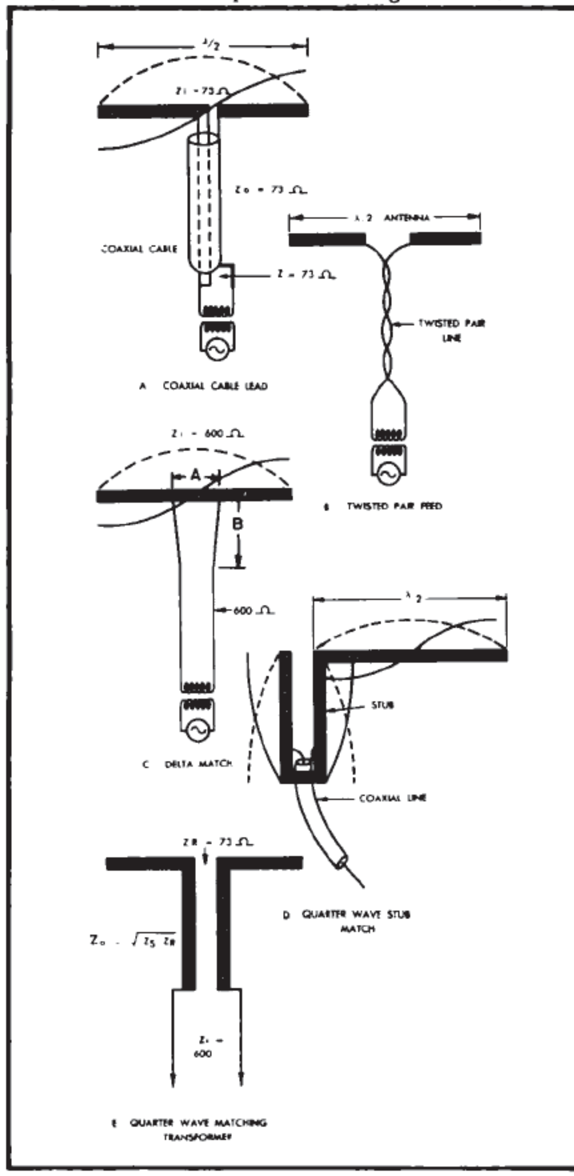


Figure 28-9 - Feeding antennas with nonresonant lines.

A1. The current and voltage are 90 degrees out of phase.

A2. The values of E and I, since impedance (Z) = E/I.

---

used as an untuned line for low frequencies. Due to excessive losses occurring in the insulation, the twisted pair is not used at higher frequencies. The characteristic impedance of such lines is about 70 ohms.

When a line does not match the impedance of the antenna, it is necessary to use special impedance-matching devices. Any of the impedance-matching lines discussed in Chapter 27 are adequate for this purpose. An example of a type of impedance-matching device is the DELTA match shown in Figure 29-9C.

Due to construction difficulties, the open, two-wire transmission line does not have a characteristic impedance (Z_0) sufficiently low to match a center-fed antenna. Practical values of Z_0 for such lines lie in a range from 400 to 700 ohms. To provide the required impedance match, a delta section (shown in Figure 28-9C) is used. This match is obtained by spreading the transmission line as it approaches the antenna. In the example given, the characteristic impedance of the line is 600 ohms and the center impedance of the antenna is 73 ohms. As the end of the transmission line is spread, its characteristic impedance increases. Proceding from the center of the antenna to either end, a point will be reached where the antenna impedance equals the impedance at the output terminals of the delta section. The delta section is then connected at this distance to either side of the antenna center.

The delta section becomes part of the antenna and, consequently, introduces radiation loss (one of its disadvantages). Another disadvantage is that trial-and-error methods are usually required to determine the dimensions of the "A" and "B" section. Since both the distance between the delta output terminals (its width) and the length of the delta section are variable, adjustment of the delta match is difficult. Its primary advantage is that it permits the use of a balanced transmission line, resulting in minimum line radiation, an essential characteristic where several antennas and lines are placed in close proximity.

Another method of matching impedance is the quarter-wave stub match shown in Figure 28-9D. In the quarter-wave stub match, the high impedance at the end of the antenna matches the open end of the stub. The impedance on the stub varies from zero at the short circuit to several thousand ohms at the open end. This makes it possible to connect a 70 ohm coaxial line a short distance from the shorted end at the 70 ohm point. It is possible to match almost any impedance along the length of the stub.

Still another impedance matching device is the quarter-wave transformer or matching transformer as shown in Figure 28-9E. This device is used to match the low impedance of the antenna to the line of higher impedance.

To determine the characteristic impedance (Z_0) of the quarter-wave section, the following formula is used:

$$Z_o = \sqrt{Z_s Z_r} \qquad (28\text{-}1)$$

where:

Z_o = characteristic impedance of the matching line.

Z_s = impedance of the feed line.

Z_r = impedance of the radiating element

For the example shown, Z_0 has a value slightly over 191 ohms. With this matching device, standing waves will exist on the antenna and $\lambda/4$ section, but not on the 600 ohm line.

Nonresonant lines are characterized by small radiation, low voltage at all points for any given power, and non-critical length. Although either resonant or nonresonant lines of one wavelength or less may be used with equal efficiency, the nonresonant lines are preferred for the transfer of power over distances greater than one wavelength.

Q3. What is the relationship between a center-fed and a current-fed one-half wave antenna?

Q4. Why is it seldom possible to connect a generator directly to an antenna?

28-4. Electrical Versus Physical Length of Dipole

The reason the types of antennas previously discussed are called dipoles should be clear. They are so named because their ends contain two equal and opposite charges.

If an antenna is constructed of very thin wire, and is isolated in space, its electrical length corresponds closely to its physical length. In practice, however, an antenna is never isolated completely from surrounding objects. For example, the antenna will be supported by insulators with a dielectric constant greater than one. The dielectric constant of air is arbitrarily assigned a numerical value equal to one. Therefore, the velocity of a wave along

a conductor is always slightly less than the velocity of the same wave in free space, and the physical length of the antenna is less (by about 5 percent) than the corresponding wavelength in space. The physical length, L, in feet, of a half-wave antenna for a given frequency (in megacycles) is derived as follows:

$$X = \frac{300}{f_{Mc}} \quad \text{and,} \quad \frac{X}{2} = \frac{300}{2f_{Mc}}$$

$$L = \frac{300 \times 3.28 \times 0.95}{2f_{Mc}} = \frac{468}{f_{Mc}} \qquad (28\text{-}2)$$

where:

f = frequency in megacycles
3.28 = feet equal to one meter
0.95 = ratio of velocity in antenna to velocity in free space

Equation 28-2 does not apply to antennas longer than one-half wavelength. If it is desired to find the length of an antenna greater than a half wavelength, the factor (0.95) will have to be introduced into the basic wavelength equation.

If an antenna is of the desired length, that is, if the electrical length approximately equals the physical length; the antenna will act like a resonant circuit, and will present to the source an impedance that will be a pure resistance. If the antenna is not of the proper length, the source will see an opposition other than the pure resistance offered under perfect conditions. The source may see an impedance that will look like a capacitive circuit or an inductive circuit depending on whether the antenna is shorter or longer than the specified wavelength. A Hertz antenna slightly longer than a half wavelength will act like an inductive circuit, and an antenna slightly shorter than a half wavelength will appear to the source as a capacitive circuit. Any two-wire open line longer than a quarter-wavelength, appears electrically as a quarter-wave section with an additional section of open-circuited transmission line attached to it. The open section, which is capacitive in itself will have its characteristics inverted and appear to the source as an inductive circuit. Compensation for the additional length can be made by cutting the antenna down to proper length, or by tuning out the inductive reactance by adding a capacitance in series. This added X_C will completely cancel the inductive reactance, and the source will then see a pure resistance, providing the proper size capacitor is used.

If an antenna is shorter than the required length, the source end of the line will appear capacitive. The reason for this is the phase relationship between the current and the voltage at this shortened open end line. The current leads the voltage by nearly 90 degrees (a condition caused only by capacitance). Because of the reflection from the open end, the generator sees a capacitive load. This condition may be corrected by adding inductance in series with the antenna. Schematic representations of antennas having proper length, longer than normal length, and shorter than normal length are shown in Figure 28-10.

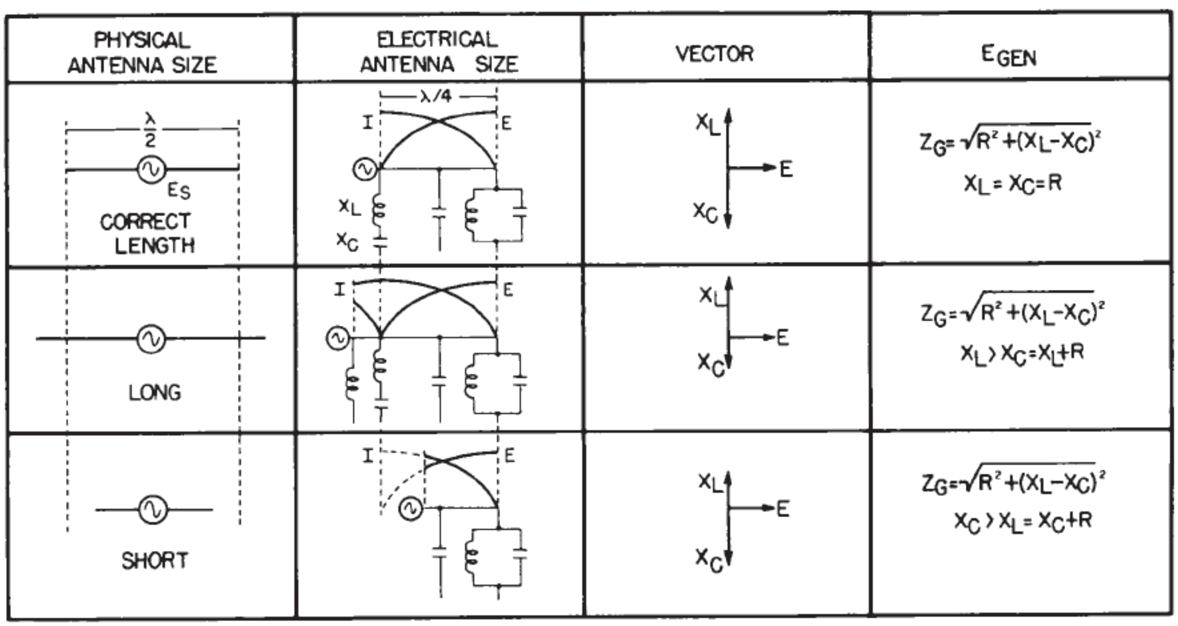


Figure 28-10 - Antenna impedance is a function of antenna length.

A3. They are the same type of feed.

A4. Usually there is a difference in the impedance of the two devices.

Q5. Why will an antenna slightly longer than a half-wave appear inductive?

28-5. Propagating a Wave from a Half-Wave Antenna

In section 28-2, it was mentioned that a quarter-wave section of transmission line, which normally is not an efficient radiating element, can be arranged to become a fairly efficient antenna. This was accomplished by splitting the quarter-wave section in half, and thereby constructing a half-wave antenna. If a voltage is applied to the line causing current to flow, the current will be maximum at the source end, or sending end, and minimum at the open end, or receiving end. The voltage however, will be maximum at the receiving end, and minimum at the sending end. The distribution of the magnetic and electrostatic fields around the half-wave antenna is shown in Figure 28-11. The magnetic field is still the most intense near the sending end, but the electric field, now extends from the positive charges at one end, to the negative charges at the other end of the line. There is no longer the cancellation of the magnetic fields, as was present in the two-wire line. As the applied voltage varies sinusoidally, both of the fields will correspondingly vary sinusoidally in intensity.

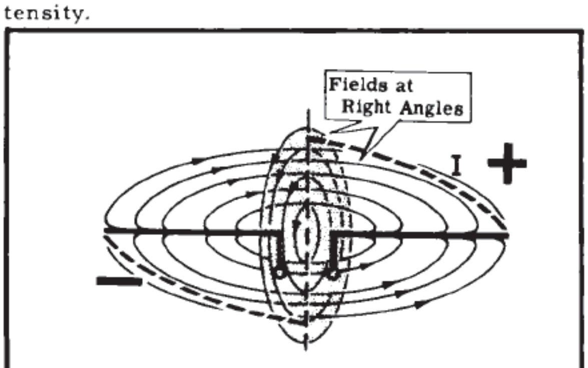


Figure 28-11 - Field distribution around a half-wave antenna.

The magnetic field is called an H field, and the electric field is called an E field. Since the current and voltage which produce these E and H fields are 90° out of phase, the fields will also be 90° out of phase, as shown by Figure 28-12.

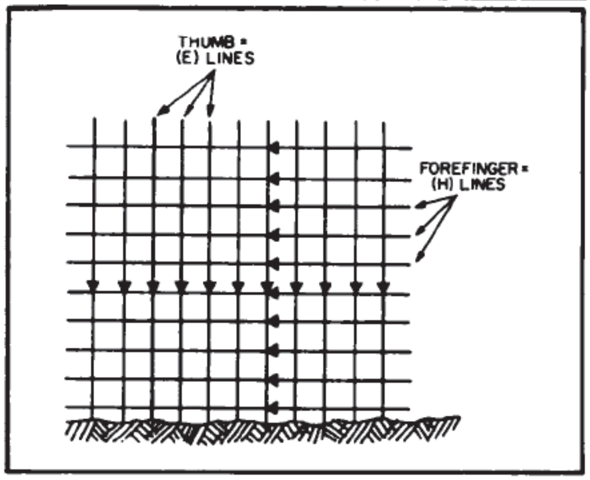


Figure 28-12 - Instantaneous E and H fields produced by antenna shown in Figure 28-11.

The manner in which energy is propagated into free space is a source of great dispute among those people concerned with it. Of all the many theories available, the following theory, has won a wide acceptance among technicians. This theory, although not exhausting all possibilities, adequately explains the phenomena. There are two fundamental fields associated with every antenna, an INDUCTION FIELD, and a RADIATION FIELD. The field that is associated with the energy stored in the antenna, is the induction field. This field is said to provide no part in the transmission of electromagnetic energy through free space. However, without the presence of the induction field, there would be no energy radiated.

To understand the manner in which the induction field is produced, a low frequency generator will be connected to an antenna, and one cycle of operation explained. Initially, let us consider the generator output to be zero, so that no fields exist about the antenna.

Assume, that the generator now produces a slight potential, having the instantaneous polarity shown in part A, of Figure 28-13. Acting as a transmission line, the antenna exhibits both capacitive and inductive properties. With the initial application of a generator potential, the antenna capacitance acts as a short, allowing a large flow of antenna current to flow in the direction indicated. This, in turn, produces a large magnetic field about the antenna. Since current is minimum at the ends of the antenna, the corresponding magnetic field at the ends will also be minimum. As time passes, charges accumulate at the ends of the antenna, which oppose antenna current and produce an electrostatic field. Eventually, the antenna capacitance

becomes fully charged and stops antenna current flow. Under this condition, the electrostatic field is maximum, and the magnetic field is fully collapsed, as shown in Figure 28-13B. As the generators potential decreases back to zero, discharging of the antenna potential occurs. During the discharging process, the electrostatic field collapses and current flows in the manner shown in Figure 28-13C. With the flow

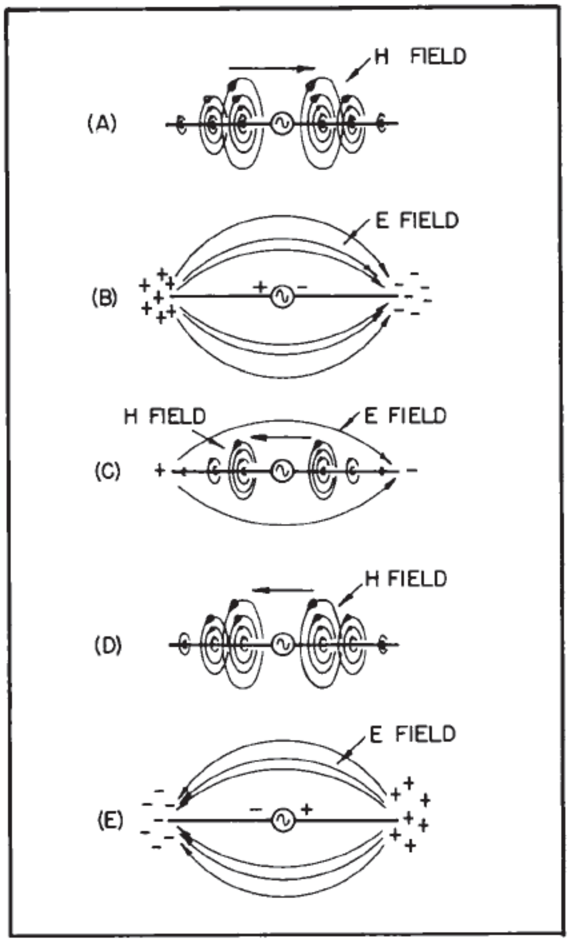


Figure 28-13 - Induction field about a half-wave dipole.

of current, an associated magnetic field is generated. Eventually, the electrostatic field completely collapses, the generator potential reverses and current is maximum, as shown in Figure 28-13D. As charges accumulate at the ends of the antenna, an electrostatic field is produced and current flow decreases. This causes the magnetic field to begin collapsing, producing more current flow, with a greater accumulated charge and electrostatic field. Eventually, the condition shown in Figure 28-13E occurs, where current is zero and the accumulated charges are maximum.

As the generator potential decreases toward zero, the antenna begins to discharge and the electrostatic field collapses. When the instantaneous generator potential becomes zero, discharge current is maximum, and the associated magnetic field is maximum. Shortly thereafter, generator potential reverses and the condition shown in Figure 28-13A recurs.

If the intensities of the magnetic (H), and electrostatic (E) fields in the vicinity of the antenna were plotted against time, a graph such as the one shown in Figure 28-14, would result. Note, that the two fields are 90° out of phase with each other. Figure 28-13, shows that the two fields are displaced 90° from each other in space (the H field existing in a plane perpendicular to the antenna, while the E field exists in a plane parallel with the antenna).

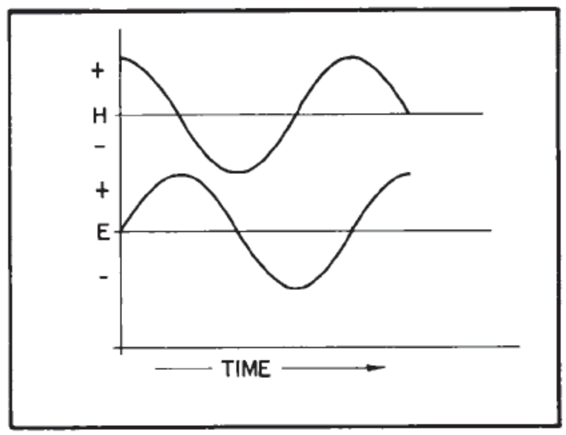


Figure 28-14 - Phase relationship of induction field components.

Since no energy was radiated from the antenna, all the energy supplied to the induction field, was returned to the antenna by the collapsing E and H fields. The intensity of the induction field decreases, as the square of the distance from the antenna. Consequently, for all practical purposes, it is considered a local field.

If the antenna generator frequency is increased, a radiation field as well as an induction field is produced. To understand the manner in which radiation occurs, it must be kept in mind, that the collapse of the E and H field involved in the induction field require time.

Although nothing has been said about the characteristics of a half-wave antenna, it is convenient to use this element in describing the mechanism of radiation. Simply stated, a half-wave antenna is one which is approximately a half-wave long at the operating frequency. For example, at 30 megacycles, a half-wave antenna is approximately 16 feet long. When power is

A5. In a line or antenna terminated in an open circuit, there is a voltage reflection from the open, that is in phase. The current reflection is 180° out of phase. Initially at the open, the voltage leads the current by 90°. If the current reflected back is done so by shifting 180°, then the current will not be leading the voltage by 90°, a characteristic of a capacitive circuit.

delivered to such an antenna, two fields are set up by the fluctuating energy: one the induction field, which is associated with the stored energy, and the other the radiation field, which moves out into space at nearly the speed of light. At the antenna, the intensities of these fields are proportional to the amount of power delivered to the antenna. At a short distance from the antenna, and beyond, only the radiation field prevails. This radiation field is made up of an electric component and a magnetic component at right angles to each other in space and varying together in intensity.

Figure 28-15 shows the manner in which the radiation field is propagated away from the antenna. The electric and magnetic field components are represented here by separate sets of flux lines, which are at right angles to each other and to the radial direction of propagation. The magnetic flux lines are shown as circular lines, having the axis of the antenna as their axis, so that they appear in the illustration as dots and crosses. The electric flux lines are closed, or endless, lines in the present case, and consist essentially of arcs of circles lying in planes containing the antenna and joined in the manner shown. These electric flux lines reverse direction, at precisely the places where the magnetic flux lines reverse. Their density varies along the radial direction in the way that the magnetic flux density varies. In fact, the electric flux density is everywhere, proportional to the magnetic flux density.

As time passes, these flux lines expand radially with the velocity of light, and new flux lines are created at the antenna to replace those that travel outward.

With a high-frequency generator connected to the antenna, the induction field is produced in the same manner as previously described. However, the generator potential reverses before the electrostatic field has had time to completely collapse. The reversed generator potential neutralizes the remaining antenna charges, leaving a resultant E field in space. This process is illustrated in Figure 28-15.

Figure 28-15A, shows a condition existing

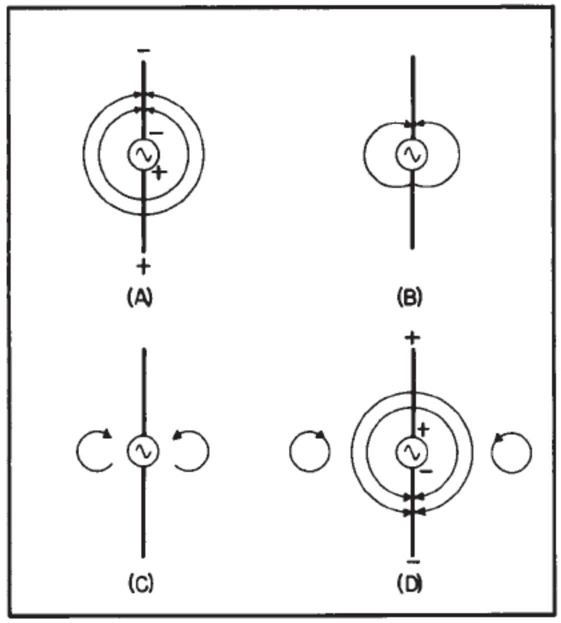


Figure 28-15 - Radiation from a half-wave antenna.

when the E field is maximum. As time passes, the generator potential decreases rapidly to zero. The E field collapses less rapidly, however, as shown in Figure 28-15B. When the antenna charges have become completely neutralized, the remaining electrostatic lines detach themselves from the antenna and form loops in space (Figure 28-15C), with no voltage to support them. When the next E field is developed, the original field is pushed away in the manner shown in Figure 28-15D.

Since each successive E field is generated with a polarity opposite that of the preceding E field (that is, the lines of force are opposite), an oscillating electric field is produced along the path of travel. When an electric field oscillates, a magnetic field is produced, having an intensity which varies directly with that of the E field. The variations in E field intensity is equivalent to a current even though it is not associated with a charge. This is called DISPLACEMENT CURRENT, and produces a magnetic field in the same manner as conduction current. The variations in magnetic field intensity, in turn, produce an E field. Thus, the two varying fields sustain each other, thereby resulting in electromagnetic wave propagation. Due to the action of the E and H components, the radiation field contains an E field, and an H field, which are in phase with respect to time but physically displaced 90° in space. This is made between

he induction field, and the radiation field.

Thus, the varying magnetic field in turn produces a varying electric field, and the varying lectric field in turn, through its associated displacement current, sustains the varying magnetic ield. Each field supports the other, and neither can be propagated by itself, without setting up he other. This is shown in Figure 28-16 where comparison is made between the induction field nd the radiation field.

The action occurring at the antenna during adiation is similar to the snapping of a whip, n that the radiation components are snapped off nd projected into space with each alternation of he applied frequency. Propagation of the radio vave occurs at nearly the speed of light; that is, 00 million meters (186,000 miles) per second.

Another factor to be mentioned about electromagnetic radiation is, that the ease with which t occurs, varies with frequency.

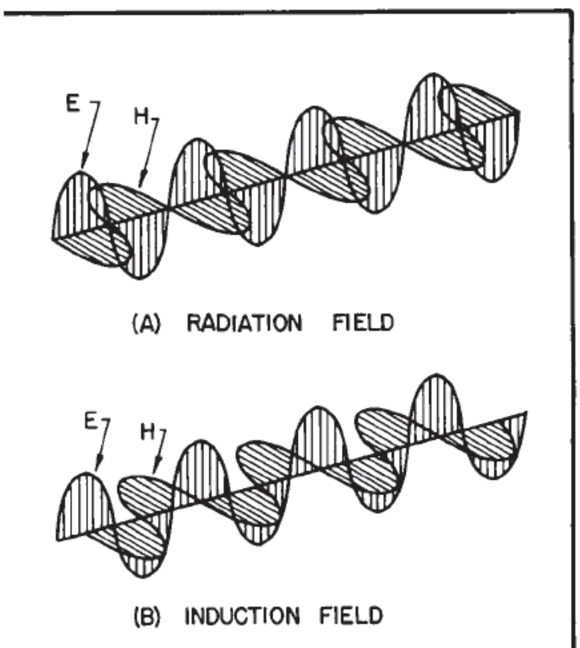


Figure 28-16 - E and H components of radiation and induction fields.

The higher the frequency, the easier the radiation becomes. At lower frequencies, such as 60 cycles per second, voltage on the antenna changes slowly, and that component of the energy radiated is so extremely small that it is of no practical value. At higher frequencies, such as 0,000 cycles per second and higher, the radiated energy is sufficient to satisfy communications requirements.

Q6. Why isn't a quarter-wave section of transmission line a good radiating element?

Q7. What is the velocity of propagation of a radiated wave in free space?

Q8. Describe an electromagnetic wave.

28-6. Polarization

In describing the principal characteristics of a wave front, the electric field component is taken as the point of reference. For example, the intensity of a radio wave usually is measured in terms of the strength of the electric field, and the orientation of the wave in space usually is described in terms of the direction of wave travel and by the direction of the electric field.

Electromagnetic fields in space are said to be polarized and the direction of the electric field is considered the direction of polarization. As the electric field is parallel to the axis of a half-wave dipole, the antenna is in the plane of polarization. When the half-wave dipole is horizontally orientated, the radiated wave is horizontally polarized. A vertically polarized wave is radiated when the antenna is erected vertically.

For maximum absorption of energy from the electromagnetic fields, it is necessary that a half-wave dipole be located in the plane of polarization. This places the conductor at right angles to the magnetic lines of force that are moving through the antenna, and parallel to the electric lines.

The polarization of a wave varies very slightly over short distances. Therefore, transmitting and receiving antennas are orientated alike, especially if short distances separate them.

Over long distances, the polarization changes. The change is usually small at low frequencies. At high frequencies, the change is quite rapid.

With radar transmissions, a received signal is a reflected wave from some object. As the polarization of the reflected signal varies with the type of object, no set position of the receiving antenna is correct for all returning signals. Generally, the receiving antenna is polarized in the same direction as the transmitting antenna.

When the transmitting antenna is close to the ground insofar as propagation is concerned, vertically polarized waves cause a greater signal strength along the earth's surface. On the other hand, antennas high above the ground should be horizontally polarized to get the greatest signal strength possible to the earth's surface. For this reason most airborne radar systems radiate horizontally polarized waves.

A6. Because of the cancellation of the magnetic fields.

A7. Electromagnetic waves are a form of radiant energy propagated into space at nearly the speed of light, 300,000,000 meters or 186,000 miles per second.

A8. A radio wave may be described as a moving electromagnetic field having velocity in the direction of travel

---

Both the induction field, and the radiation field contain electric and magnetic fields. The E and H components of the induction field are 90° out of phase IN TIME. In the radiation field, the E and H field are 90° out of phase IN SPACE, but they are in phase in time.

The position of the half-wave antenna in space is important, because it affects the POLARIZATION of the electromagnetic wave.

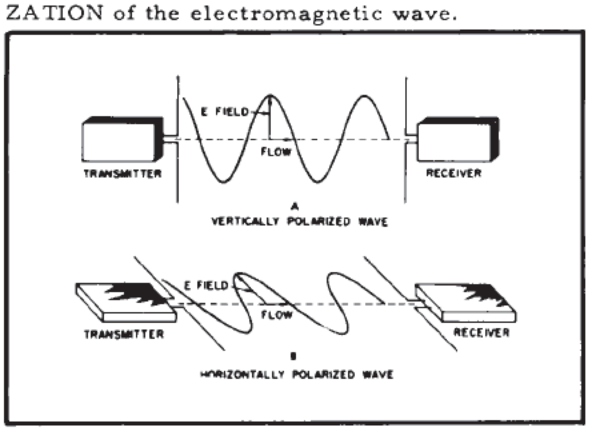


Figure 28-17 - Polarized waves.

By convention, the plane of polarization of a radio wave is the plane in which the E field propagates with respect to the earth. If the E field component of the radiated wave travels in a plane perpendicular to the earth's surface (vertical), the radiation is said to be VERTICALLY POLARIZED, as shown in Figure 28-17A. Likewise, if the E field propagates in a plane parallel to the earth's surface (horizontal), the radiation is said to be HORIZONTALLY POLARIZED, as shown in Figure 28-17B.

If the plane of polarization remains unchanged, the process is known as LINEAR polarization. Since the E field propagates radially from the antenna axis, a linearly polarized wave will have a plane of polarization which coincides with the antenna axis. That is, with linear polarization, a vertical antenna will produce vertical polarization. By various means, it is possible to obtain a linearly polarized wave at any angle.

A common method of representing the E and H fields of a wavefront is through the use of vectors, as shown in Figure 28-18. Figure 28-18A, shows a horizontally polarized wave, while Figure 28-18B, shows a vertically polarized wave. The instantaneous intensity of each field is represented by the respective vector length while the instantaneous field direction (or polarity), is shown by vector direction (using conventional rectangular coordinate designations).

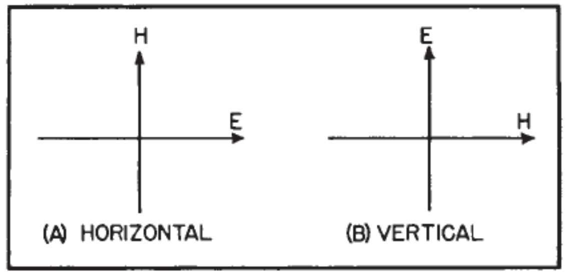


Figure 28-18 - Vector representation of radiated wave front.

Through the use of a right-hand rule, the direction of wave propagation can be determined if the directions of the E and H components are known. This rule states, that if the thumb, forefinger, and middle finger of the right hand are extended so that these three digits are mutually perpendicular, the middle finger will point in the direction of wave propagation if the thumb points in the direction of the E field and the forefinger points in the direction of the H field. In Figure 28-18, then, wavefront A would propagate toward the reader while wavefront B would propagate away from the reader. It should be noted, that since both the E and H fields reverse directions simultaneously, propagation of a particular wavefront is always in the same direction (away from the antenna).

A linearly polarized wave at an angle of 45° from the horizon is shown in Figure 28-19A. Considering only the E field vector, vertical and horizontal components of this field, can be represented in the manner that is shown in Figure 28-19B. Due to the angle of linear polarization (45°), the intensity of the vertical and horizontal components are equal. The two components are also in phase with each other, so that at any given instant, E_V and E_H have the same amplitude. If, now, either component is shifted in phase by 90° (or one-quarter wavelength), a new type of polarization is produced, called CIRCULAR POLARIZATION.

To understand the manner in which circular polarization occurs, consider an antenna tilted 45°, in the manner previously discussed. Consider further, that the horizontal component (E_H)

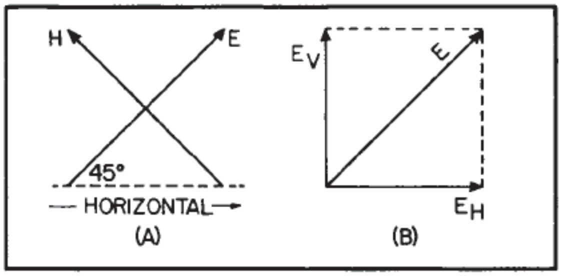


Figure 28-19 - Wavefront polarized 45° from horizontal.

of the E field is shifted 90° behind the vertical component (E_V). This condition is represented by the waveforms of Figure 28-20. Although the two waveforms are of equal peak amplitude, due to the 90° phase shift one is maximum when the other is zero. Since these waveforms represent the instantaneous magnitudes of E_V and E_H, by combining these magnitudes vectorially, the instantaneous intensity of the E field can be determined (as well as its direction).

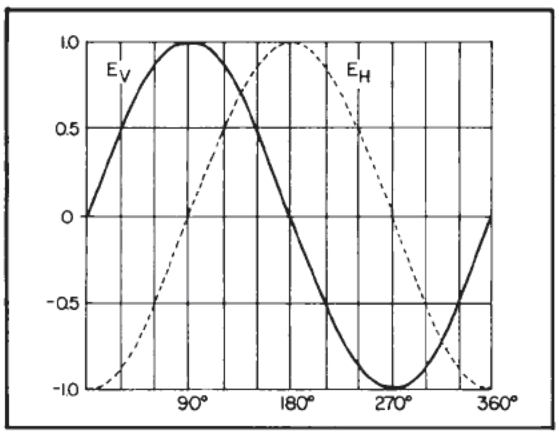


Figure 28-20 - Horizontal and vertical components of E field for circular polarization.

From Figure 28-20, the values of E_V and E_H at 30 degree intervals have been recorded in Table 28-1.

Referring to the vector diagram in Figure 28-19B, it can be seen that E_V and E_H are two sides of a right triangle while E is the hypotenuse. If the value of the two sides are known, the value of the hypotenuse can be determined by:

$$E = \sqrt{E_V{}^2 + E_H{}^2} \qquad (28\text{-}3)$$

By inserting values of E_V and E_H (obtained from Table 28-1) into this expression, the instantaneous magnitude of the E field can be determined at 30 degree intervals. For example, if E_1

Degrees	E_V	E_H
0	0.0	-1.0
30	+0.5	-0.866
60	+0.866	-0.5
90	+1.0	0.0
120	+0.866	+0.5
150	+0.5	+0.866
180	0.0	+1.0
210	-0.5	+0.866
240	-0.866	+0.5
270	-1.0	0.0
300	-0.866	-0.5
330	-0.5	-0.866
360	0.0	-1.0

Table 28-1 - Instantaneous values of E_V and E_H at 30° intervals.

represents the electric field at zero degrees, its value would be:

$$E_1 = (0.0)^2 + (-1.0)^2 = 0+1 = 1$$

To determine the direction of the E field, a vector diagram, as shown in Figure 28-21A, is used.

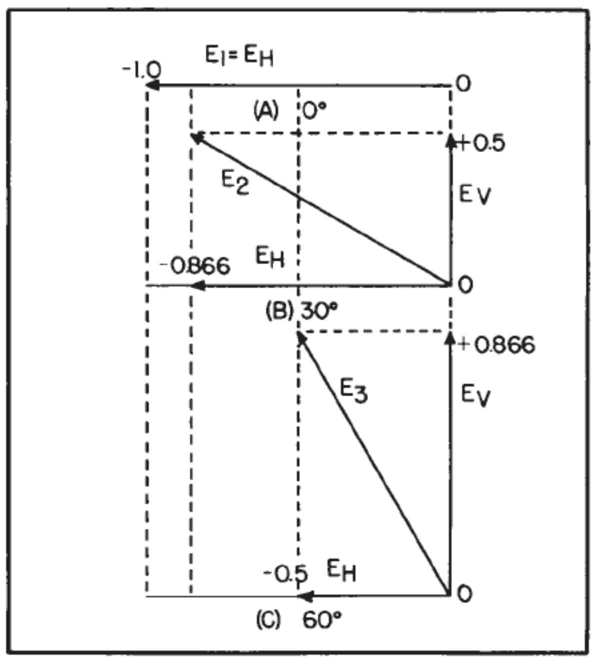


Figure 28-21 - E field vectors at 0°, 30° and 60° points of Figure 28-20.

Since there is no vertical component of the E field at 0°, the magnitude of E_1 equals that of E_H and is projected in the negative horizontal

direction.

If E_2 represents the E field intensity at 30°, E_3 at 60°, and continuing in like fashion to 360°, vector diagrams can be made to show the direction of the E field at 30° intervals. Figure 28-21B and C represent E field vectors at 30° and 60° respectively. At 30°, the value of E_2 is:

$$E_2 = \sqrt{E_V{}^2 + E_H{}^2} = \sqrt{(0.5)^2 + (-0.866)^2} = \sqrt{0.25 + 0.75}$$

$$= \sqrt{1} = 1$$

while at 60°, E_3 has a value of:

$$E_3 = \sqrt{E_V{}^2 + E_H{}^2} = \sqrt{(0.866)^2 + (-0.5)^2} = \sqrt{0.75 + 0.25}$$

$$= \sqrt{1} = 1$$

Notice that the intensity of the E field remains constant but its direction changes.

Figure 28-22 shows the E field vector at 30° intervals for one complete cycle. To an observer, the field intensity remains constant but its plane of polarization constantly rotates in a clockwise direction, thereby resulting in circular polarization. The circular motion may be either clockwise or counterclockwise, depending on which component is shifted in phase and the direction of the shift. One application of circular polarization is in radar to enable polarization through such polarization-insensitive barriers as rain, fog and cloud formations.

As previously explained, circular polarization requires that the vertical and horizontal components of the E field be 90° out-of-phase (electrically) and their peak amplitudes be equal. If the required phase displacement is produced

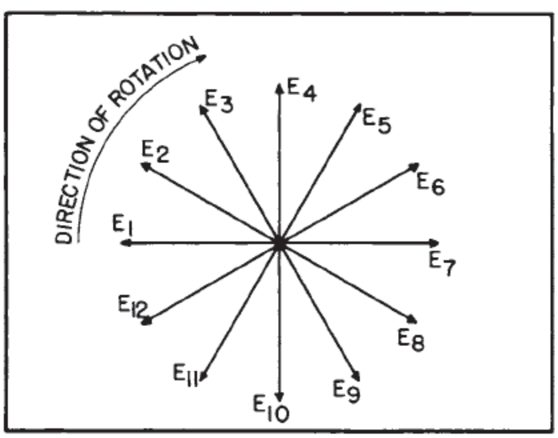


Figure 28-22 - Circular polarization.

but the peak amplitudes are not equal, ELLIPTICAL POLARIZATION occurs. Two possibilities of elliptical polarization exist, depending on which component has the greater amplitude. If the peak amplitude of the vertical component is larger, the E field intensity will vary in the manner shown in Figure 28-23A. Figure 28-23B shows elliptical polarization when the peak amplitude of the horizontal component is larger.

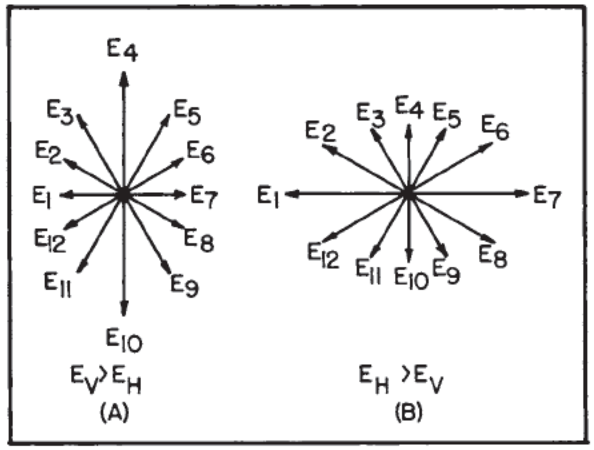


Figure 28-23 - Elliptical polarization.

The type of polarization used in a specific antenna depends upon the type of transmission desired. The following factors affect polarization and require consideration for a specific antenna installation.

1. A horizontally polarized wave of low or medium frequencies assume the characteristics of a vertically polarized wave. This is due to the shorting effect of the earth on the electric field parallel to it at these frequencies. At frequencies below 5 megacycles, the ground acts as a conductor. At higher frequencies, however, the ground acts more like a dielectric. A vertically polarized wave at these frequencies is unaffected and propagates with linear polarization.

2. At high frequencies transmitting antennas are generally horizontally polarized, but the polarization of the wave varies because it usually splits into several components which follow different paths. Consequently high-frequency receiving antennas can be polarized in any direction.

3. At extremely high frequencies transmission follows line-of-sight directions so that a single component arrives at the receiving antenna, rather than a combination of

components experienced at high frequencies. Due to this fact, the receiving antenna must be polarized in the same direction as the transmitting antenna to obtain maximum efficiency.

4. When antennas are close to the ground, vertically polarized waves yield a stronger signal close to the earth than do horizontally polarized waves. If the transmitting antenna is at least one wavelength above the ground, either type of polarization produces the same field intensity near the surface of the earth. With transmitting antennas several wavelengths above the ground, however, horizontally polarized waves produce a stronger signal close to the earth than can be achieved with vertical polarization.

Q9. Will a horizontally polarized high frequency wavefront be received by a vertically polarized antenna? Explain.

28-7. Electromagnetic Propagation from a Dipole

The relationship of the E and H fields as they are propagated in the radiation field is such that they are in phase in time and 90 degrees out of phase in space. Along the half-wave antenna, the intensity of the field is not uniform. There are points along its length where the field is at a maximum, and other points along its length where the field is minimum. This is shown in Figure 28-24.

The E field is shown as the solid, closed loops that exist on each side of the antenna. The H fields, which are 90 degrees out of phase in space, are shown as the circles that enclose either crosses or dots (to indicate direction). The sine waves that are shown superimposed on the fields indicate the variation in electric flux intensity at various distances and angles away from the antenna.

In a direction perpendicular to the antenna, the fields are strongest. In a parallel direction away from the antenna, that is, off the ends of the antenna, the fields are weakest. For this reason, the half-wave antenna is said to be a directional antenna.

Because the current is greatest at the center of a dipole, maximum radiation takes place at this point and practically no radiation takes place from the ends. If this antenna could be isolated completely in free space, the points of maximum radiation would be in a plane perpendicular to the plane of the antenna at its center. The doughnut-shaped surface pattern is shown in Figure 28-25A, and the horizontal cross section pattern is shown in Figure 28-25B. Because a circular field pattern is created, the field strength is the same in any compass direction.

Theoretically, a vertical dipole in free space has no vertical radiation along the direct line of its axis. However, it may produce a considerable amount of radiation at other angles measured to the line of the antenna axis. Figure 28-25C shows a vertical cross section of the radiation pattern of Figure 28-25A. The radiation along OA is zero; but at another angle, represented by angle AOB, there is appreciable radiation. At a greater angle, AOC, the radiation is still greater. Because of this variation in field strength pattern at different vertical angles, a field-strength pattern of a vertical half-wave antenna taken in a horizontal plane must specify the vertical angle of radiation for which the pattern applies.

Figure 28-25D, shows half of the doughnut pattern for a horizontal half-wave dipole. The maximum radiation takes place in the plane perpendicular to the axis of the antenna and crosses through its center.

The variation in radiation field intensity about an antenna can be shown graphically by polar diagrams as in Figure 28-26. Zero distance is assumed to be at the center of the chart indicating the center of the antenna and the circumference of the tangent circles is laid off in angular degrees. Computed or measured values of field strength then may be plotted radially in a manner that shows both magnitude and direction for a given distance from the antenna. Field strengths in the vertical plane are plotted on a semicircular polar chart (not shown in the figure) and are referred to as vertical polar diagrams.

Q10. Why is it said that the half-wave antenna is directional.

28-8. Radiation Resistance

An understanding of the power dissipation in the electromagnetic fields produced by an antenna may be obtained by considering the familiar concepts of power and phase angle in an ac circuit, such as a resonant tank circuit. It will be recalled that, in an ac circuit, true power is determined by:

$$P = EI \cos \theta \qquad (11\text{-}12)$$

where θ is the phase angle between current and voltage. In an ideal tank circuit, each reactive component produces a 90° phase shift between current and voltage. Since the cosine of 90° is zero, the power dissipated over a complete cycle is zero.

An antenna may be considered as a tank circuit. Since the magnetic field is directly proportional to antenna current, it may be used to compute the power dissipated. In the induction field,

A9. It may be received because the original horizontally polarized wavefront of high frequency may be shifted as it is transmitted through space, and could conceivably arrive at the receiving antenna as a vertically polarized wave.

A10. Because the intensity of the radiation field it produces varies over its length.

---

the electric and magnetic components are 90° out of phase, as previously illustrated in Figure 28-16B. Consequently, no power is dissipated by the induction field. Any power delivered to the field during one portion of a cycle is returned during another portion.

The electric and magnetic components of the radiation field are in phase, as previously illustrated by Figure 28-16A, and power is therefore dissipated. The power, which is radiated by the antenna to form the familiar radiation patterns, is comparable to the power dissipated by the resistance of a practical tank circuit. The value of resistance that would dissipate the same power that the antenna dissipates is called RADIATION RESISTANCE. Since a relationship exists between the power dissipated by the antenna and the antenna current, radiation resistance can be mathematically defined as the ratio of total power dissipated to the square of the effective value of antenna current, or:

$$R = \frac{P}{I^2} \qquad (28\text{-}4)$$

where: R = radiation resistance in ohms
I = effective value of antenna current at the feed point in amperes.
P = total power radiated from the antenna length, as shown in the graph of Figure 28-27.

The radiation resistance varies with antenna length, as shown in the graph of Figure 28-27. For a half-wave antenna, the radiation resistance measured at the current maximum (center of the antenna) is approximately 73 ohms. For a quarter-wave antenna, the radiation resistance measured at its current maximum is approximately 36.6 ohms. These are free-space values; that is, the values of radiation resist-

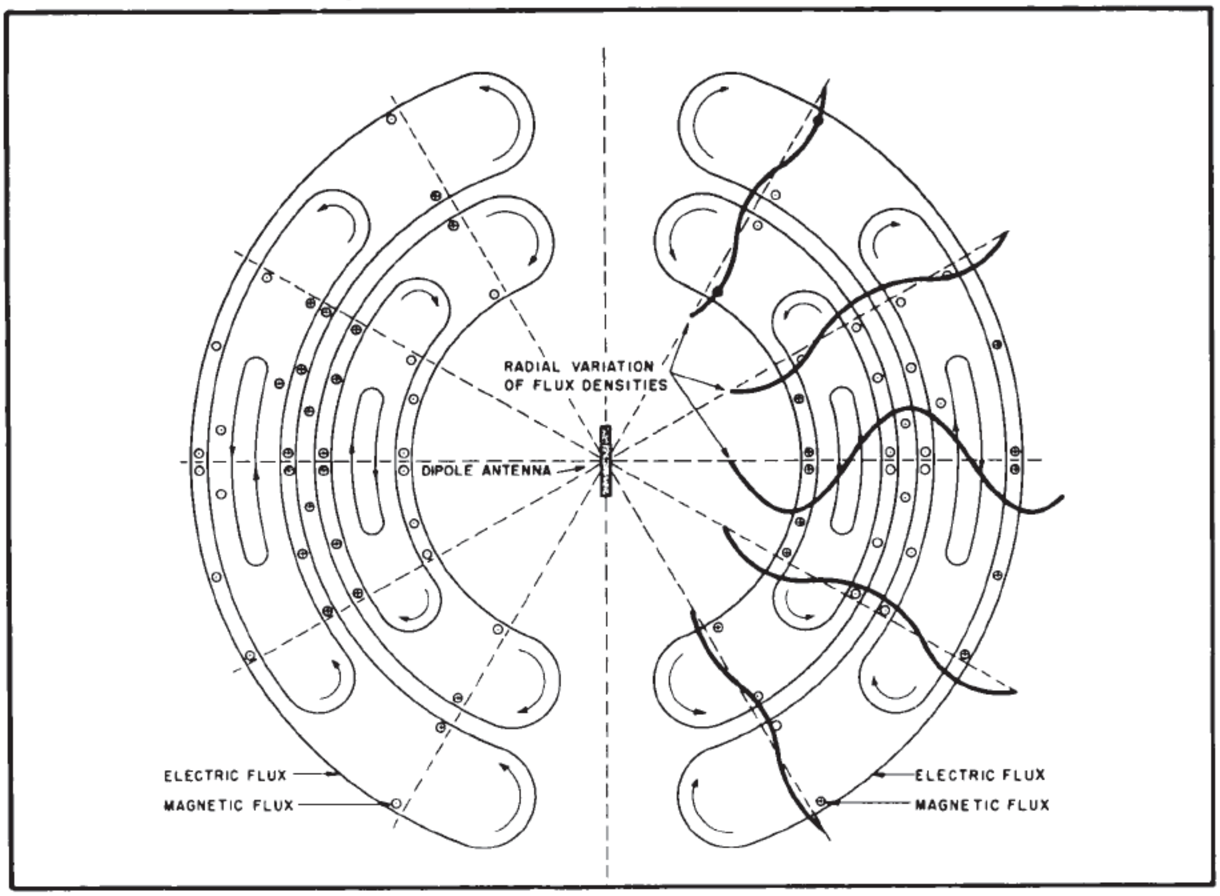


Figure 28-24 - Relationship of fields around a half-wave antenna.

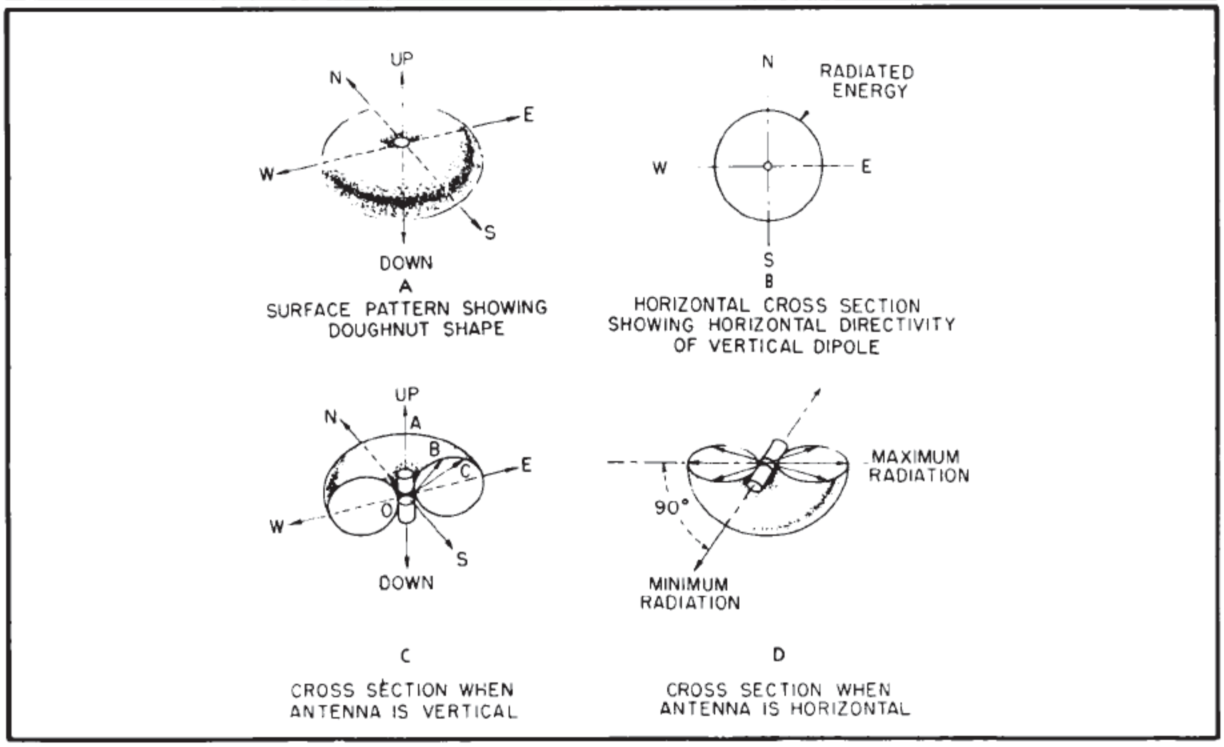


Figure 28-25 - Radiation pattern of a dipole.

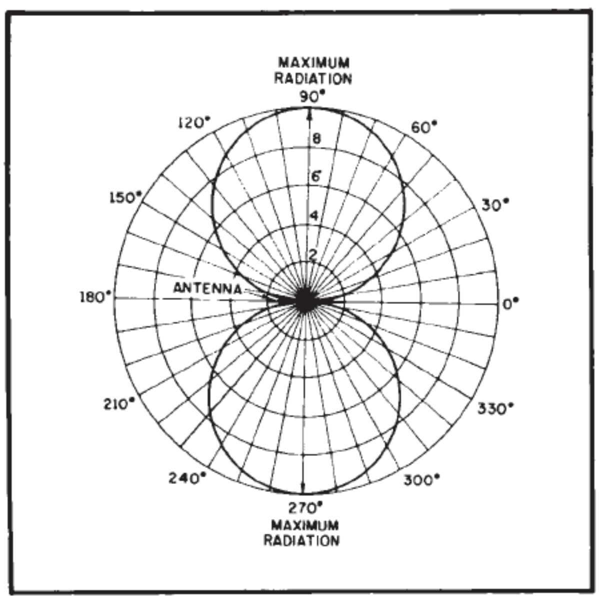


Figure 28-26 - Polar diagram of an antenna showing relative field strength.

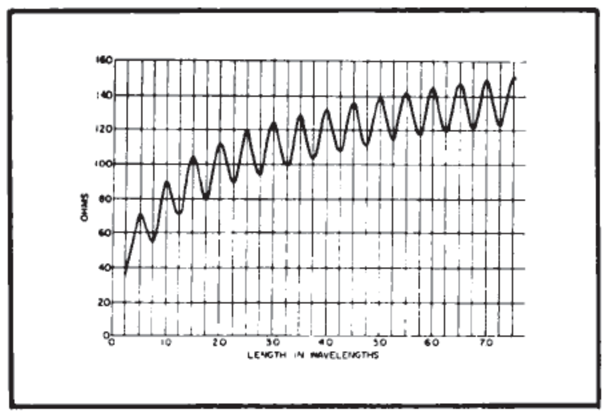


Figure 28-27 - Radiation resistance of antennas in free space plotted against length.

ance which would exist if the antenna were completely isolated so that its radiation pattern would be unobstructed.

For practical antenna installations, the height of the antenna affects radiation resistance. Changes in radiation resistance occur because of ground reflections which intercept the antenna and alter the amount of antenna current flowing. Depending on their phase, the reflected waves may increase antenna current or decrease it. The phase of the reflected waves arriving at the antenna, in turn, is a function of antenna height and orientation.

At some antenna heights, it is possible for a reflected wave to induce antenna currents in phase with transmitter current so that total

774-422 O - 64 - 8

antenna current increases. At other antenna heights, the two currents may be 180° out of phase so that total antenna current is less than if no ground reflection occurred.

With a given input power, if antenna current increases, the effect is as if radiation resistance decreases. Similarly, if the antenna height is such that the total antenna current decreases, the radiation resistance is increased. The actual change in radiation resistance of a half-wave antenna at various heights above ground is shown in Figure 28-28. The radiation resistance of the horizontal antenna rises steadily to a maximum value of 90 ohms at a height of about three-eighths wavelengths. Then the radiation resistance falls steadily to 58 ohms at a height of about five-eights wavelengths. The resistance then continues to rise and fall around an average value of 73 ohms, which is the free-space value. As the height is increased, the amount of variation keeps decreasing.

The variation in radiation resistance of a vertical antenna is much less than that of the horizontal antenna. The radiation resistance is a maximum value of 100 ohms when the center of the antenna is a quarter-wavelength above ground. The value falls steadily to a minimum value of 70 ohms at a height of a half-wavelength above ground. The value then rises and falls by several ohms about an average value slightly above the free-space value of a horizontal half-wave antenna.

Since antenna current is affected by antenna height, the field intensity produced by a given antenna also changes. In general, as the radiation resistance is reduced, the field intensity increases; whereas an increase in radiation resistance produces a drop in radiated field intensity.

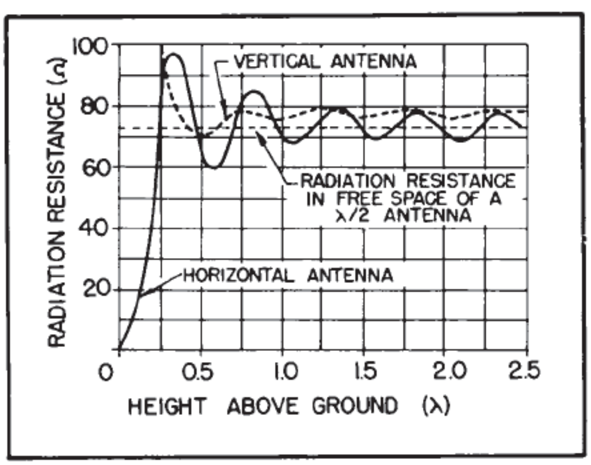


Figure 28-28 - Radiation resistance of half-wavelength antennas at various heights.

28-9. Antenna Impedance

Since there is energy stored in the induction field about the antenna, the presence of the reactive inductance and capacitance components is evident. The value of these reactances and the value of the radiation resistance will affect the value of current flow in the antenna. The combination of a reactive and a resistive opposition renders some impedance value for the antenna. This antenna impedance is similar to the characteristic impedance of a transmission line, and is called the ANTENNA INPUT IMPEDANCE. The formulas that may be used to compute the value of this input impedance are as follows:

$$Z = \frac{E}{I} \qquad (12\text{-}5)$$

or:

$$Z = R \pm jX \qquad (11\text{-}7)$$

Any antenna at resonance presents a specific impedance at every point along its length. This can be seen by comparing the voltage and current values distributed along an antenna as shown in Figure 28-29A. Using the Ohm's law impedance formula, it can be seen that the highest impedance occurs where current is lowest and vice versa. Between points of highest and lowest impedance, antenna impedance values follow the curve of Figure 28-29B. As previously mentioned, the impedance at the center of a half-wave antenna is approximately 73 ohms, whereas the high impedance points at the ends are approximately 2500 ohms.

The voltage and current distribution along an antenna in free space depends upon whether it is resonant or nonresonant at the applied frequency. Since it is impossible to completely isolate an antenna from ground, surrounding objects, etc., the voltage and current are changed by the inductive and capacitive effects introduced. This, in turn, changes the impedance values along the length of the antenna and must be considered when determining the method of coupling power to the antenna.

In the impedance formula ($Z = R \pm jX$), R represents total antenna resistance which is

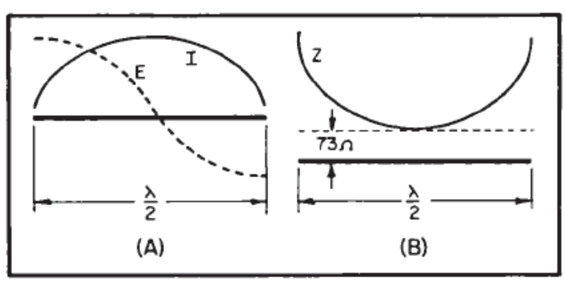


Figure 28-29 - E, I and Z distribution for a half-wave antenna.

composed of three individual resistances to be discussed later. The reactive component (jX) has a value if the antenna is not resonant. The half-wave antenna is the shortest resonant length of antenna. However, antennas which are 2 or more half-wavelengths may be resonant. Such antennas are said to operate on harmonics. If an antenna is 4 half-wavelengths of the transmitter frequency, it is being operated at the fourth harmonic of its lowest resonant frequency. In other words, this antenna is a half-wavelength at 1/4 the frequency of operation.

The diagram in Figure 28-30 is a graph of the input resistance of center-fed antennas for various wavelengths. Resistance values for both a thin and a thick antenna are plotted so that the effect of the diameter of the radiating element is apparent. In Figure 28-31, the reactance is plotted as a function of wavelength.

The curves show that the antenna may be either inductive or capacitive depending on its length, and that abrupt changes of impedance occur at multiples of a half-wavelength. The points in Figure 28-31 where the reactance

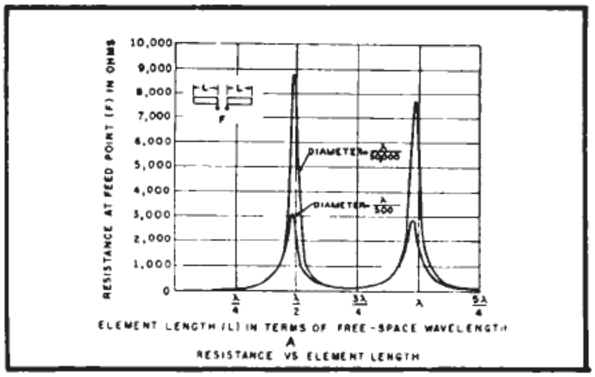


Figure 28-30 - Resistance versus antenna length.

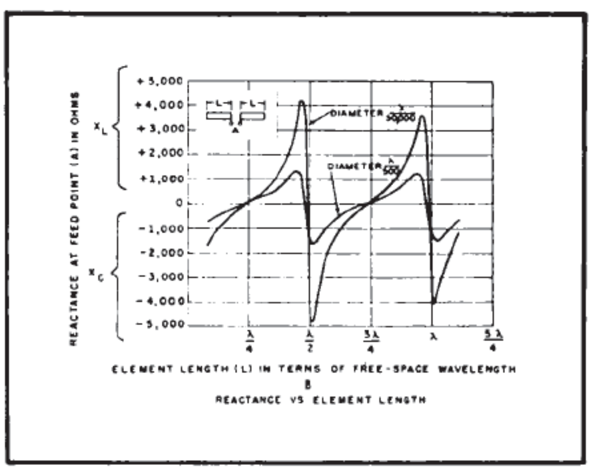


Figure 28-31 - Reactance versus antenna length.

curves cross the zero point indicate the resonant lengths of the antenna. Because the curves are plotted in terms of the free-space wavelength, the effect of the reduced velocity of the wave motion along the antenna is shown by the curves. For example, a half-wave antenna element is resonant only when it is less than the free-space half wavelength. The foreshortening is caused by the increased capacitance associated with the elements. If the diameter of the radiator is large, for example λ/500, the increased capacitance is greater than for the thin element. As a result, the large-diameter radiator is foreshortened more than the thin radiator.

Figure 28-31 may be used to calculate the input impedance of a center-fed antenna. For example, let it be required to find the impedance of a thin (diameter = λ/50,000) antenna having a half-wavelength equal to five-eighths of the wavelength being fed to it. In this application the antenna is not fully resonant. The impedance includes both resistance and reactance. The resistance is located on the curve in Figure 38-31 halfway between λ/2 and 3λ/4, and is approximately 150 ohms. Similarly, the reactance is found on the diagram in Figure 28-31 to be capacitive with a value of approximately 1,100 ohms. The input impedance at the 5λ/8 point is therefore:

$$Z = 150 - j\,1,100 = 1,100\,\underline{/-82.2^\circ}$$

Therefore, for maximum transfer of energy to the antenna, a feedline to a 5λ/8 center-fed antenna in free space must be designed to present an impedance of $1,100\,\underline{/-82.2^\circ}$ ohms. In this case, the feedline has a resistance of 150 ohms and an inductive reactance of 1,100 ohms.

The current flowing in an antenna is opposed by three kinds of resistance. With the antenna considered as a radiator of energy, the power dissipated in the form of radiation can be considered an I^2R_r loss where R_r is the radiation resistance. As a conductor, the antenna dissipates power in the form of heat. This power is an I^2R_o loss where R_o is the ohmic resistance of the antenna. Another I^2R loss (called dielectric loss) is due to leakage resistance of dielectric elements, such as insulators. This value of resistance is usually included in the ohmic resistance.

The function of a transmitting antenna is to dissipate as much energy as possible in the form of radiation. The energy dissipated by the radiation resistance, therefore, is the useful part of the total power dissipated. Since actual power loss depends on ohmic resistance (including leakage resistance), this resistance should be kept as low as possible. In the half-wave antenna, the radiation resistance is large com-

pared to the ohmic resistance, and most of the available energy is radiated.

Unless the ground behaves as a nearly perfect conductor, the amplitude of the ground-reflected wave will be much less than the amplitude of the wave before reflection. A portion of the wave is absorbed by the ground and constitutes GROUND LOSSES. As can be seen from Table 28-2, the conductivity of various surfaces varies greatly.

Ground Material	Relative Conductivity
Sea Water	4, 500
Flat, rich soil	15
Average flat soil	7
Fresh water lakes	6
Rocky hills	2
Dry, sandy, flat soil	2
City residential area	2
City industrial area	1

Table 28-2 - Relative conductivity of various surfaces.

From Table 28-2 it can be seen that salt water provides the best transmitting surface.

One manner frequently used to increase the conductivity of the ground under an antenna is through the use of a GROUND SCREEN. This screen consists of a fairly large area of metal mesh laid on or directly beneath the surface of the ground under the antenna. Sometimes the metal screen is laid on a wooden framework that is erected 8 to 12 feet off the ground with wires running from the mesh to grounding rods embedded in the earth. In practice, the screen extends about half a wavelength in all directions, although larger surfaces produce better conductivity. Two advantages in using ground screens is a reduction in ground losses and more accurate predictions of radiation patterns (since the radiation resistance can be precalculated).

The input impedance of an antenna is affected by the presence of nearby conductors (for example: rigging on ships, guy wires, etc.). Any object that can be affected by the induction field will distort the field, and also the antenna current and voltage distribution. Therefore, the input impedance of the antenna will be changed, and necessary corrections must be made to provide the best match to each antenna. Because this effect is almost always difficult if not impossible to calculate, corrections are usually made using the trial-and-error method.

Q11. If either the reactance or the radiation resistance increases, what happens to the input impedance of the half-wave antenna? Explain.

OTHER BASIC ELEMENTS

28-10. The Marconi Antenna

It was mentioned in section 28-2 that the Marconi antenna is used primarily with frequencies below 2 mc. The difference between the Marconi antenna and the Hertz antenna is that the Marconi type requires a conducting path to ground, and the Hertz type does not. The Marconi antenna is a quarter-wave grounded antenna or any odd multiple of a quarter wavelength.

A Marconi antenna used as a transmitting element is shown in Figure 28-32. The transmitter is connected between the antenna and ground. The actual length of the antenna is one-quarter wavelength. However, this type of antenna, by virtue of its connection to ground, uses the ground as the other quarter wavelength, making the antenna electrically a half-wavelength. This is so because the earth is considered to be a good conductor. In fact, there is a reflection from the earth which is equivalent to the reflection that would be realized if another quarter-wave section were used. By use of the Marconi antenna, which is a quarter-wave in actual physical length, half-wave operation may be obtained. All of the voltage, current, and impedance relationships characteristic of a half-wave antenna will also exist in this antenna. The only exception to this is the input impedance which is approximately 36.6 ohms.

One of the most frequent uses for the Marconi antenna is in portable equipment. In naval vessels, the additional quarter wavelength required for half-wave operation is usually provided by the ship's hull.

The effective current in the Marconi grounded antenna is maximum at the base and minimum at the top while voltage is minimum at the bottom and maximum at the top.

When the conductivity of the soil in which the Marconi antenna is supported is very low, the reflected wave from the ground may be greatly attenuated. The reflection from the ground is made in much the same way as a voltage wavefront is reflected from a resonant transmission line. A great attenuation of the reflected signal is highly undesirable. To overcome this disadvantage, the site location could be moved to a location where the soil possesses a high conductivity. If it is impractical to move the site, provisions must be made to improve the reflecting characteristics of the ground.

When an actual ground connection cannot be used because of the high resistance of the soil or a large buried ground screen is impractical, a COUNTERPOISE may replace the usual direct ground connection. The counterpoise consist of a structure made of wire erected a short dis-

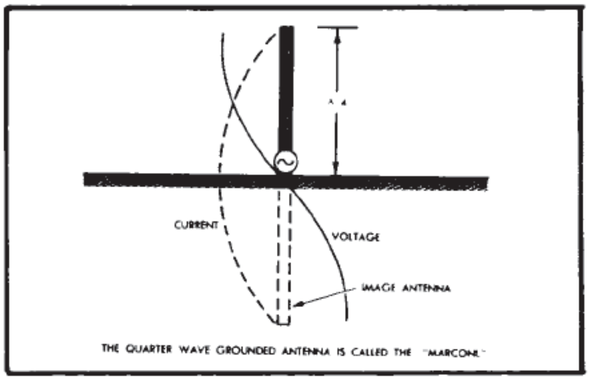


Figure 28-32 - Grounded Marconi antenna.

tance above the ground and INSULATED FROM THE GROUND. The size of the counterpoise should be at least equal to and preferably larger than the size of the antenna.

The counterpoise and the surface of the ground form a large capacitor. Due to this capacitance, antenna current is collected in the form of charge and discharge currents. The end of the antenna normally connected to ground is connected through the large capacitance formed by the counterpoise. If the counterpoise is not well insulated from ground, the effect is much the same as that of a leaky capacitor with a resultant loss greater than if no counterpoise is used.

Although the shape and size of the counterpoise are not particularly critical, it should extend for equal distances in all directions. Contact between counterpoise and antenna is usually made from 8 to 12 feet above ground. When the antenna is mounted vertically, the counterpoise may have any simple geometric pattern, such as is shown in Figure 28-33. The counterpoise is constructed so that it is nonresonant at the operating frequency. The operation realized by use of either the well grounded Marconi antenna, or the Marconi antenna using counterpoise, is the same as that of the half-wave antenna of the same polarization.

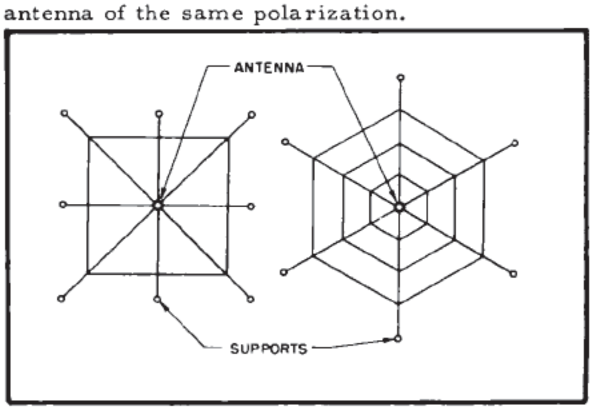


Figure 28-33 - Counterpoise (top view)

Q12. How must a grounded Marconi antenna be polarized. Why?

28-11. Bent Antenna

A widely used method for increasing the effective height of the current loop in a vertical antenna is through the use of a BENT ANTENNA, as shown in Figure 28-34. Such an antenna is commonly referred to as an INVERTED L or FLAT-TOP antenna. This design is particularly useful at low frequencies where a vertical half-wave antenna becomes impractically high.

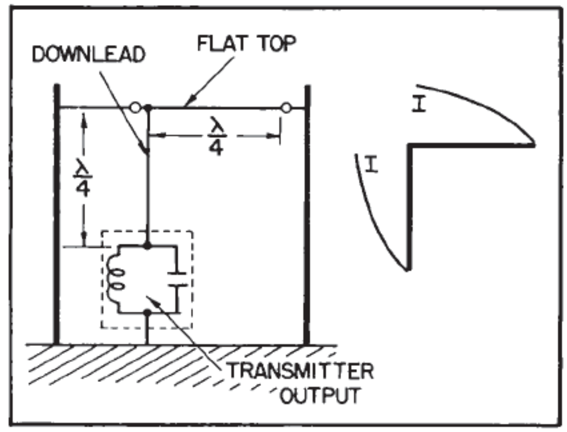


Figure 28-34 - Inverted L antenna and current distribution.

In Figure 28-34 the vertical portion of the antenna, called a downlead, is a quarter-wavelength and connects to the transmitter. The horizontal portion, called a flat top, is also a quarter-wavelength and connects to the downlead to form the second half of the antenna. As can be seen by the current distribution of the antenna, maximum current occurs at the junction of the vertical and horizontal sections of the antenna. This results in greater efficiency than could be achieved by a vertical quarter-wave antenna but with the same physical height. However, considerable radiation leaves the antenna at a high angle, in a direction opposite the free end. Such high-angle radiation is undesirable at low frequencies where propagation should be directed along the earth's surface.

Improvement in the ground-wave propagation characteristics of the inverted L antenna can be achieved through the use of multiple flat tops connected to a single downlead. Even better results are achieved by folding the flat top of an inverted L antenna into a u-shape, thus the vertical component of the radiated field can be minimized. With such an antenna, practically all radiation occurs in a horizontal plane from the downlead, resulting in good ground-wave transmission.

A11. The impedance also increases because the impedance is directly proportional to the value of the radiation resistance and reactance.

A12. The grounded Marconi antenna is polarized vertically because it is close to the ground, an area which provides greater signal strength if the antenna is vertically polarized.

28-12. Long-Wire Antenna

A LONG-WIRE antenna is one that has distributed over its length two or more half-waves of energy. The long-wire antenna is also referred to as a HARMONIC ANTENNA.

To analyze the characteristics of a long-wire antenna, it is first necessary to review some of the basic functions of the half-wave antenna. This is done because the long-wire operation is similar to the operation of the half-wave antenna.

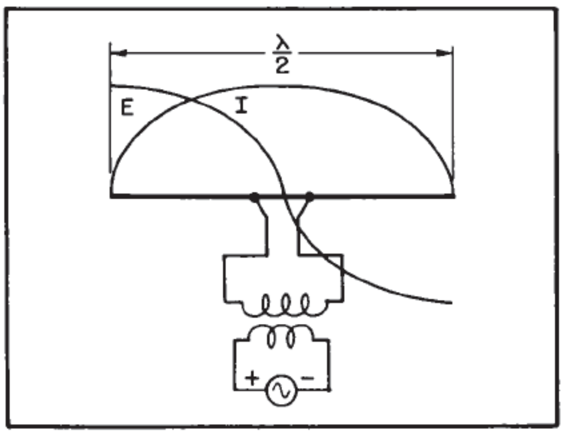


Figure 28-35 - Center-fed half-wave antenna.

Because of the polarities indicated on the generator in Figure 28-35, current in the left-hand side of the antenna flows toward the generator. In the right-hand side of the antenna, current flows away from the generator. The result is a current wave similar to the one shown above the half-wave antenna. Since it is a half-wave antenna, the current at the end is minimum, and the voltage at the end is maximum.

If the same center-fed line is increased by a half-wavelength, as shown in Figure 28-36, the antenna will now be a two half-wave antenna. The only difference between this antenna and the one shown in Figure 28-35 is that this antenna can accomodate two half-waves of current. Electrically, this configuration is called a DRIVEN COLLINEAR ARRAY. The term "array" means a group of elements which, considered

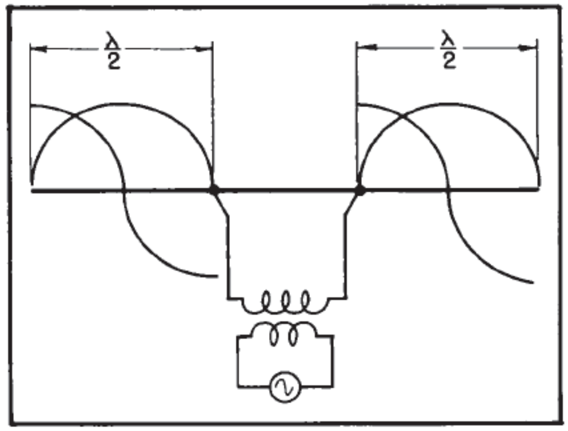


Figure 28-36 - Two half-wave antennas.

separately, could be individual antennas but are combined to produce a given radiation pattern. In this case, the array is nothing more than two half-wave antennas operating in phase at their fundamental frequency. Although this type of configuration partially satisfies the definition of a long-wire antenna, it is not, strictly speaking, a long-wire antenna since current does not reverse in adjacent half-wave sections. However, this antenna may be easily converted into a long-wire antenna by simply moving the generator under a current lobe or to the end of the line. The long-wire antenna is shown in Figure 28-37.

With the generator connected as shown in Figure 28-37A, current flows from left to right in the half-wave section of the antenna. The direction of current flow in the other section is reversed. If the generator is moved to the end of the line, the current distribution on the line will be as shown in Figure 28-37B. It is still a long-wire antenna. If the antenna is any odd number of half wavelengths (1 1/2, 3 1/2, etc.) so that a current loop occurs at the center of the

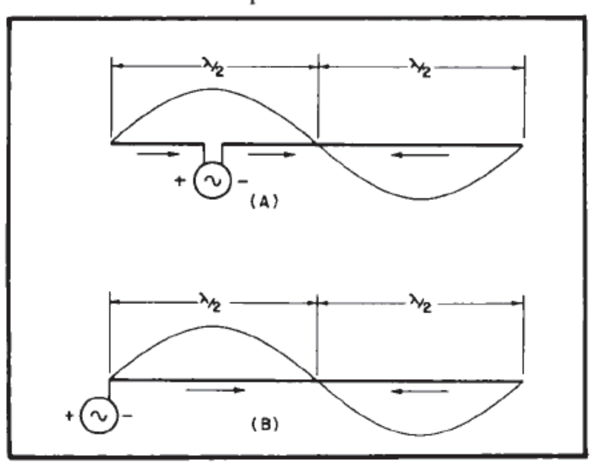


Figure 28-37 - Long-wire antenna.

antenna, center feeding may be used. A long-wire antenna may be considered to be one which is made up of a number of half-wave sections fed 180 degrees out of phase and separated a half-wavelength apart.

As might be expected, the change in length has a considerable effect on the radiation intensity and the RADIATION PATTERN. A radiation pattern is a valuable tool to aid the technician in determining which type of antenna is suitable for a particular application. Figure 28-38 shows the radiation pattern that may be expected from antennas of various lengths.

The cloverleaf patterns indicate the intensity of the field in their respective positions. These cloverleafs are lobes. Those areas in which no lobes appear are the nulls. Notice that as the length of the antenna increases, radiation occurs more in line with the direction of the antenna; that is, there is a pronounced radiation increase toward the ends of the antenna. This is a marked advantage as compared to the basic half-wave antenna. Of course, radiation still occurs at right angles to the antenna. Consequently, the resultant maximum radiation occurs at some acute angle with respect to the wire, and not completely at right angles or completely along the line of the wire.

The radiation pattern of a long-wire antenna is such that it contains one pair of lobes for each of its half-wavelength sections. Two pairs of these lobes will be major, the remaining will be minor. Figure 28-38B shows an antenna whose radiation pattern contains four lobes (2 pairs) and it consists of two half-wavelength sections. Its total length is therefore equal to one wavelength. The three pairs of lobes in the radiation pattern of Figure 28-38C, indicates that the antenna length equals 1 1/2λ.

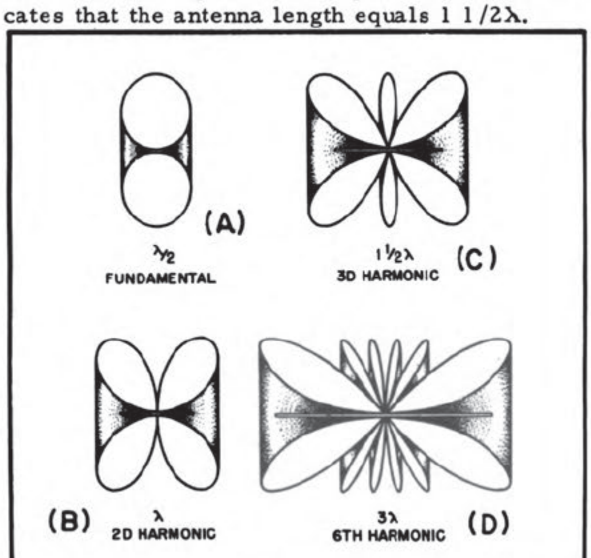


Figure 28-38 - Typical radiation patterns.

The long-wire antenna possesses a greater gain than that of the half-wave antenna. It possesses a wider angle of radiation, and, as the length of the line increases, there are more minor lobes produced.

Q13. What is meant by the term "array"?

Q14. What may be gained by adding more segments to a long-wire antenna?

28-13. Antenna Gain

The term GAIN has an application in antenna discussions. However, it is not specifically the same type of gain that is normally encountered in a discussion concerning amplifier stages.

The basic half-wave and quarter-wave antennas previously discussed, whether polarized horizontally or vertically, are directive antennas. It was pointed out that these antennas will provide greatest radiation in a direction perpendicular to the antenna, and least radiation at the ends of the antenna. The directive characteristics not only apply to transmitting antennas but apply to receiving antennas as well.

All antennas have some directional characteristics even though many of them are classified as having the ability to transmit and receive with equal ability in all directions. An ideal situation would be to have a standard of measure so that the degree of directivity of different antennas may be compared. This is an impractical situation. However, an imaginary standard antenna will serve the purpose. This idealized antenna is called an ISOTROPIC ANTENNA. An isotropic antenna is one that radiates and receives equally in all directions; that is, it is omnidirectional. If power is applied to the isotropic radiator, it will radiate electromagnetic energy in all directions. Some of this radiated energy can be detected and recorded at some distant receiving station. If the same amount of power is applied to a half-wave antenna under exactly the same conditions, this antenna will produce a field strength greater in some directions than that produced by the isotropic radiator. When the half-wave antenna produces a field in one direction which is greater in intensity than the field produced in other directions, it is usually at the expense of the other field. It is thus said that a half-wave antenna produces a GAIN in field strength in a direction at right angles to itself, and a loss in field strength in other directions. In other words, there will be a gain in a half-wave antenna broadside to it, and a loss at the ends.

Antenna gain is rated in terms of decibels. For the half-wave antenna, the field strength or the power ratio may be converted into decibels

A13. It means any configuration of antenna elements.

A14. A greater field intensity towards the ends of the conductor.

---

as follows:

Gain in db = 20 log X field strength ratio (28-5)

Gain in db = 10 log X power ratio (28-6)

Since there is no such thing as an isotropic radiator, a half-wave antenna is used as a standard.

Q15. Why is gain important?

28-14. Antenna Tuning Circuit

Energy flow from a transmitter to an antenna is accomplished by various types of coupling networks. One such network, commonly used with shipboard communications, is shown in Figure 28-39.

The antenna tuning system includes the antenna coupling capacitor C_1, the antenna tuning inductor L_1, the antenna tuning capacitor C_2, and the antenna feed switch S_1. The dc blocking capacitor, C_3, is connected in series with C_1, to protect the antenna system from dc potentials that might cause damage or voltage breakdown of C_1. The antenna tuning capacitor, C_2, is variable and is operated in one of two circuit arrangements. With S_1 in position S, capacitor C_2 is connected in series with L_1, and the antenna is current (or series) fed. In position P, S_1 connects L_1 in parallel with C_2, and the antenna is voltage fed.

The antenna system is tuned by first adjusting C_1 to minimum coupling, and tuning the final power amplifier stage to resonance. Then, capacitor C_2 and inductor L_1 are tuned for antenna resonance.

The antenna now appears as a pure resistance to the final amplifier. The capacitance of C_1 is then increased in small steps until the required loading of the final amplifier is obtained. Each time the coupling capacitor is changed, however, the final amplifier and antenna tuning circuit (L_1C_2) must be returned to resonance. Care should be exercised to prevent overcouple with C_1. After the final amplifier and the antenna circuit are resonated, the load on the final amplifier is purely resistive, and maximum transfer of energy from the final amplifier to the antenna is obtained.

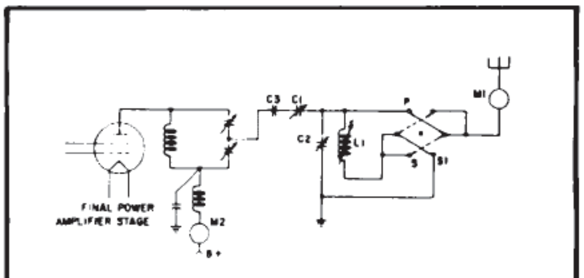


Figure 28-39 - Antenna coupling to transmitter.

Q16. What is the purpose of antenna tuning?

WAVE PROPAGATION

After an electromagnetic wave is transmitted from a transmitting antenna, it must be propagated through space to a receiving antenna thereby establishing a useful communications system. This section is devoted entirely to an explanation of how electromagnetic energy travels through free space.

Since the medium of travel for electromagnetic energy may be free space, it is necessary that the nature of free space be known so that its effects on the quality of transmission may be predicted. Weather conditions, changes in the level of radiation from the sun, and physical obstructions on the earth's surface all affect the quality and reliability of transmission. Because weather conditions and changes in the radiation level from the sun are not, as yet, controlled by man, any adverse conditions caused by their variation must simply be tolerated until a superior wireless communication system not subject to these variations can be developed.

28-15. Radio Frequency Spectrum

In section 1-7, the electromagnetic spectrum was shown. Between the audio spectrum and the infrared spectrum, there is a range of frequencies known as the RADIO FREQUENCY SPECTRUM. This spectrum is subdivided and classified in Table 3.

Description	Frequency (mc)	Abbreviation
Very low frequency	0.01 to 0.03	VLF
Low frequency	0.03 to 0.3	LF
Medium frequency	0.3 to 3	MF
High frequency	3 to 30	HF
Very high frequency	30 to 300	VHF
Ultra high frequency	300 to 3,000	UHF
Super high frequency	3,000 to 30,000	SHF

Table 28-3 - Radio-frequency spectrum.

28-16. Components of the Propagated Spectrum

Primarily, there are two types of transmitted electromagnetic waves, the GROUND WAVE and the SKY WAVE. The ground waves are those that travel near the surface of the earth. These waves are greatly affected by the conductivity of the earth and any obstruction such as mountains or buildings on its surface. Ground wave transmission is used primarily in local communications.

The ground wave is composed of three waves: the SPACE or TROPOSPHERIC WAVE, the EARTH REFLECTED WAVE, and the SURFACE WAVE. The space wave is also called the DIRECT WAVE.

The sky wave is an electromagnetic wave which is propagated at such an angle that it travels up through the atmosphere, strikes its upper layer, the ionosphere, and refracted back toward the earth. Sky wave transmission is used in long distance transmission.

28-17. The Surface Wave

The SURFACE WAVE is that part of the ground wave which travels in contact with the earth's surface. Because of the conductivity of the earth's surface, some of the energy of the surface wave will be absorbed by the ground. Since the earth's surface in most locations is an excellent conductor, the passing electromagnetic energy will cause eddy currents to flow in the ground. These eddy currents dissipate power and are classified as a loss. Eddy current losses are greatest when the surface wave is polarized horizontally and least when the surface wave is polarized vertically. The surface wave and the rate at which it diminishes is shown in Figure 28-40.

Since the electrical properties of the earth along which the surface wave travels are relatively constant, the signal strength from a given station at a given point is fairly constant. This holds true in nearly all localities except those that have a distinct rainy or dry season. Changes in the amount of moisture causes the conductivity of the soil to change. If the conductivity is high, the absorption will be low and transmission using the surface wave can be achieved over a considerable distance. If the conductivity is low, the absorption of the transmitted energy will be high and the distance of transmission realized by use of the surface wave will be short.

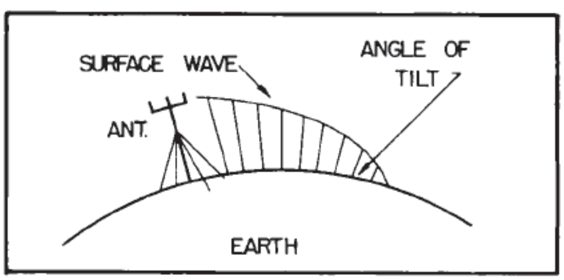


Figure 28-40 - The surface wave.

Different types of terrain will have various effects on the absorption of radio frequencies in the surface wave. If the surface wave is transmitted over water, especially salt water, transmitting distance of the surface wave is greatly increased as illustrated in Figure 28-41.

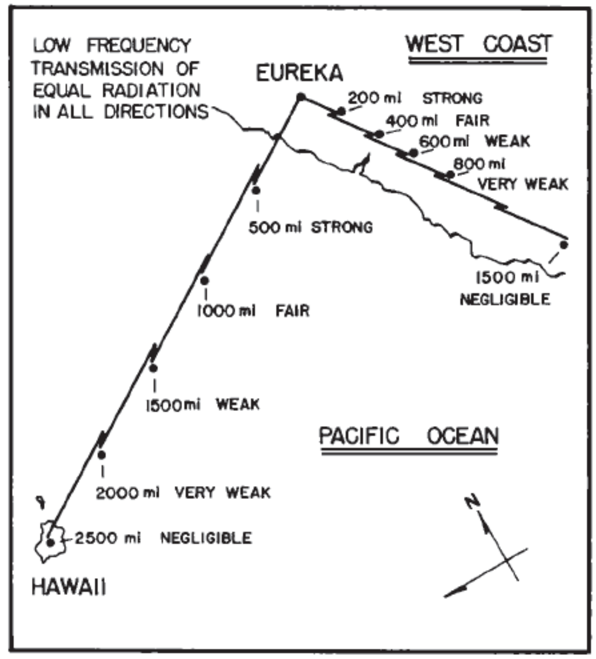


Figure 28-41 - Effectiveness of ground wave propagation over land and sea water.

Figure 28-40 shows the surface wave being tilted toward the ground. The angle made between the plane of transmission and the surface of the earth is known as the ANGLE OF TILT. If the conductivity of the earth's surface is high, the angle of tilt is large and little energy is absorbed by the ground. On the other hand, if the conductivity of the earth is low, the angle of tilt is low and the absorption of energy is high.

Shore-based stations utilize a high-power surface wave, transmitted, preferably, over water, to provide long-range, surface-wave communications over appreciable distances.

Q17. What effect does polarization have on the surface wave?

28-18. Direct and Earth-Reflected Waves (Space Waves)

The DIRECT and EARTH-REFLECTED WAVES are shown in Figure 28-42. They are also called SPACE WAVES or TROPOSPHERIC WAVES. The name tropospheric wave is given

A15. It represents the degree of antenna directivity.

A16. To insure maximum transfer of power.

A17. Polarization affects the rate of absorption of the electromagnetic wavefront. The horizontally polarized wave is absorbed more than the vertically polarized wave.

---

since the medium through which the space wave travels is the troposphere, that portion of the atmosphere directly above the earth's surface. The troposphere will be described in greater detail in topic 28-19.

As shown in the diagram, the direct wave is that component of the space wave that travels in almost a straight line from the transmitting antenna to the receiving antenna. This type of transmission, strictly utilizing the direct wave, is known as LINE-OF-SIGHT TRANSMISSION. Line-of-sight transmission means that both the transmitting and receiving antennas are optically visible to one another.

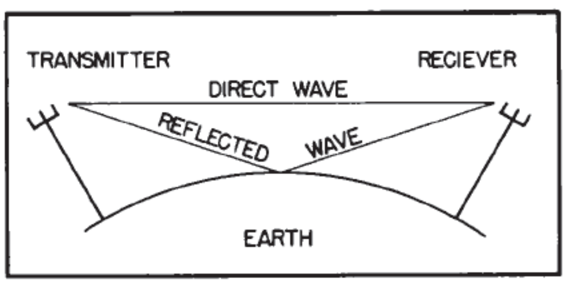


Figure 28-42 - Direct and ground reflected waves (space waves).

It was mentioned that the direct wave travels in almost a straight line. The direct wave is slightly bent by tropospheric REFRACTION. This causes the direct wave to be bent back toward the earth, and extends transmission beyond the optical horizon.

Refraction of electromagnetic energy occurs in a manner similar to the refraction of light. Refraction of electromagnetic or light energy is caused when the energy passes from one density medium to another.

An example of refraction is the apparent bending of a spoon when it is immersed in a vessel of water. The bending seems to take place at the surface of the water, or exactly at the point where there is a change of density. Obviously, the spoon does not bend from the pressure of the water. The light forming the image of the spoon is bent as it passes from the water, a medium of high density, to the air, a medium of comparatively low density.

The bending of light is shown in Figure 28-43. Also shown is the reflected wave, the wave which is reflected from the surface of the water. Notice that the angle of incidence is equal to the angle of reflection. If the direction

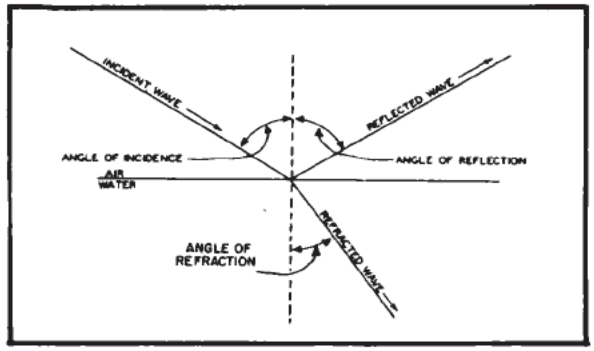


Figure 28-43 - Bending of a wavefront of light by refraction.

of travel were reversed, that is from water to air, the bending would be the same.

At frequencies greater than 30 mc, the attenuation of the surface wave is so extensive that communications is primarily by use of the direct wave. However, the ground reflected wave may cause a decrease in the intensity of the direct wave when they both arrive at the receiving antenna.

The ground reflected wave is that portion of the space wave which is reflected from the surface of the earth. The intensity with which the wave is reflected is dependent upon the COEFFICIENT OF REFLECTION of the surface which it strikes, and the angle of incidence. The relationship between the direct wave and the ground reflected wave was shown in Figures 28-42 and 28-43. To further understand the action of this reflection, consider the diagram in Figure 28-44. This diagram shows that the reflection of light and the reflection of electromagnetic energy occurs in much the same way. Although the angle of incidence is equal to the angle of reflection, there is a change in the phase of the incident and reflected wave, as seen by

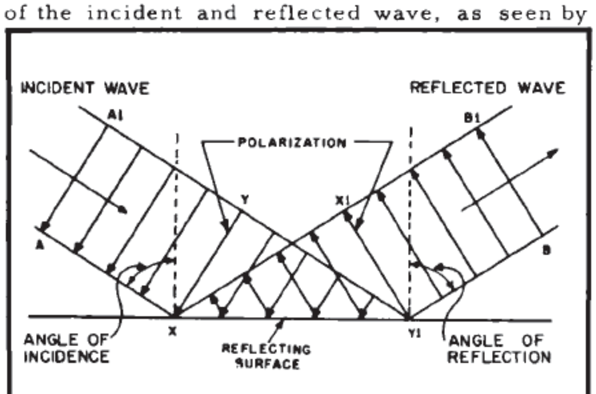


Figure 28-44 - Reflection of a wavefront.

the difference in the direction of polarization. The incident and reflected waves are 180 degrees out of phase.

Different material will have a different coefficient of reflection The coefficient of reflection could also be called the coefficient of absorption because the intensity of the reflected wave is dependent upon the quantity of incident wave energy absorbed by the surface of reflection. Some materials will absorb more energy than others. Energy not absorbed will be reflected in much the same way as a voltage wavefront from a shorted transmission line.

The ground-reflected wave begins approximately at that point and angle where the surface wave ceases to have appreciable effect. Since the ground reflected wave exists between the surface wave and the direct wave, there are only a few degrees where reflection from the ground occurs. As the angle of transmission reaches an angle almost parallel to the surface of the earth, reflection reduces to a minimum.

The ground-reflected wave is usually undesirable, for between points lower than a few thousand feet, and separated by a few miles, it can cause cancellation of the direct wave at the receiving antenna. The incident wave is in phase with the direct wave. The reflected wave however, is 180 degrees out of phase with the incident wave. If both the reflected and the direct waves arrive at the receiving antenna simultaneously, some cancellation will occur. The cancellation is partial because the reflected wave takes a longer time to make the trip from the transmitting antenna to the ground and then to the receiving antenna than does the direct wave, which travels almost .n a straight line from the transmitting antenna to the receiving antenna. Therefore, the phase difference is not exactly 180 degrees. Also, the comparative intensity with which the direct and ground reflected waves arrive at the receiving antenna is different because the intensity of the ground reflected wave is reduced due to partial absorption of the incident wave by the ground.

The relationship between all three waves is shown in Figure 28-45. Only a few lines representing these wavefronts are shown. However, the number of actual waves moving in all directions from the transmitting antenna to the receiving antenna is very large. Also shown in the diagram are some waves that have no effect on the transmission of energy over the distance indicated, but are present nevertheless.

Thus far, a communication system utilizing line-of-sight transmission has been discussed. Suppose that an obstruction exists between the transmitting antenna and the receiving antenna. If the height of the obstruction is slightly higher than the receiving antenna, the receiving antenna will not be optically visible from the transmitting antenna. The communication system would appear to be broken; however, it is not. As with light waves, the electromagnetic waves have the ability to bend as they pass the edge of an obstruction. This bending phenomena, quite different from refraction, is called DIFFRACTION.

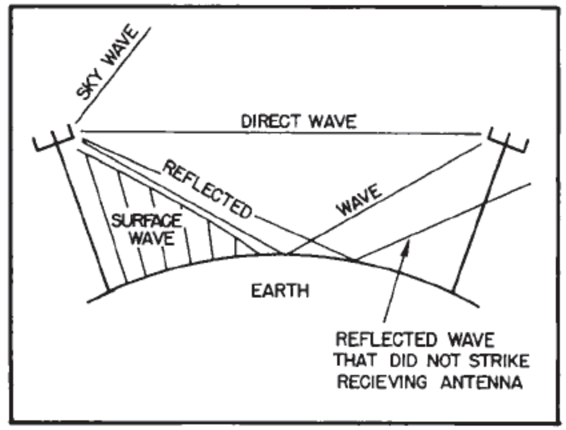


Figure 28-45 - Complete ground wave with sky wave.

Electromagnetic waves in the radio frequency spectrum bend around an object to a greater extent than do light waves. In fact, the lower the frequency of the energy, the greater the bending. This is shown in Figure 28-46. The zone or area which is obstructed, where electromagnetic waves are not present, is called the SHADOW ZONE. .The earth itself will cause the diffraction of electromagnetic waves causing them to follow the contour of the earth. The diagram in Figure 28-46 shows that electromagnetic waves are diffracted over the top and around the side of an obstruction.

Therefore, diffraction and refraction can cause the transmission distance to be greater than the line-of-sight distance. Even if the

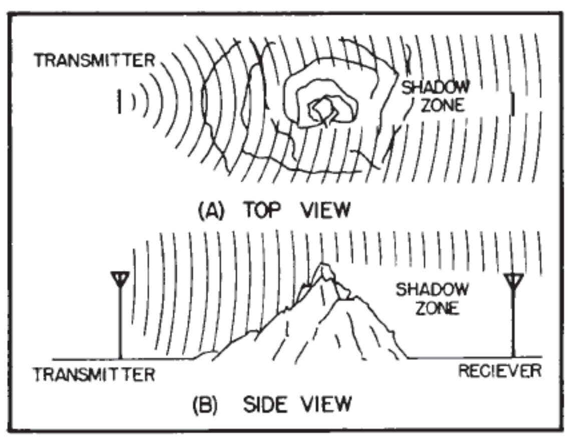


Figure 28-46 - Diffraction around an object.

receiving antenna is beyond the optical horizon, or if it is located behind an obstruction, transmission may still be received. A new term called RADIO HORIZON is now appropriate and refers to the actual distance between two antennas by which transmission may occur. Of course, there is a limit to the extended range realized through refraction and diffraction. A receiving antenna could be so completely hidden, or be so far beyond horizon, that transmission by use of the surface and space waves would be impossible. A general equation to determine the radio horizon distance in miles is:

$$D = \sqrt{2H} \qquad (28\text{-}7)$$

where: D = distance of radio horizon in miles
H = height of transmitting antenna in feet

Notice that in the equation (28-7) the height of the receiving antenna is not taken into account. The actual distance from the transmitting antenna to the receiving antenna is equal to the sum of the horizon distances for each individual antenna. An equation expressing this distance is:

$$D = \sqrt{2H_1} + \sqrt{2H_2} \qquad (28\text{-}8)$$

where: D = distance of the radio horizon in miles
H_1 = height in feet of transmitting antenna in feet
H_2 = height of receiving antenna in feet

It was mentioned that refraction occurs when a wave passes from a medium of heavy density to a medium of comparatively light density. At times, there can be such turbulence in the atmosphere when normal refractions will be exaggerated. One of these radical atmospheric changes is called TEMPERATURE INVERSION. Temperature inversion may be caused by several events. A warm air mass overrunning a cold mass, the rapid cooling of the surface air after sunset, the sinking of an air mass heated by compression, and the heating of air above a cloud layer by reflection of the sun's rays from the upper surface of the clouds all cause temperature inversions. These temperature inversions, causing radical refractions occur frequently and may cause signal FADING at the receiving station; or at least a diminished field intensity. Fading is any fluctuation in the intensity of a received signal and will be considered in greater detail later in this chapter. It should be noted that neither tropospheric refraction, nor diffraction cause a phase shift, whereas, reflection from a surface does.

Q18. What is the difference between refraction and diffraction?

Q19. What is the difference between optical horizon and radio horizon?

Q20. What is the phase relationship between the incident wave and the refracted wave?

28-19. Troposheric Duct

The troposphere is the layer of the atmosphere which lies directly above the surface of the earth. It extends from the surface to approximately an altitude of 6.5 miles. Ordinarily it is an area in which the density is high near the earth's surface and becomes progressively lower as altitude increases.

Propagation characteristics in the troposphere vary under special weather conditions; and in some areas remain fixed for an extended period of time. In the tropics and at sea, temperature inversions are present almost continuously at altitudes up to 3,000 feet. When the boundary of the temperature inversion is sharply defined, electromagnetic waves traveling horizontally, or at low angles of elevation with respect to the earth, are trapped by the refracting layer of air and continue to be bent back toward the earth. However, the refracted wave does not travel back and strike the earth. It is refracted upward by the other boundary of the temperature inversion, and it continues this action in a back-and-forth fashion. It is trapped. The trapping medium is called a DUCT· specifically, the TROPOSPHERIC DUCT. The waves contained in this duct may be trapped for considerable distances, or until the temperature inversion is normalized. Therefore, waves traoped in the duct may be transmitted over great distances which at times extend over thousands of miles. It should be remembered that the troposphere extends only 6.5 miles high. This means that the tropospheric duct may be only a few thousand feet high. To realize communications by use of this duct, both the transmitting and receiving antennas should be in the duct. If not, the waves will be guided away from the antenna in much the same way as waves are guided along a transmission line. A diagram showing the tropospheric duct is illustrated in Figure 28-47.

To be trapped in the tropospheric duct, the wave must be transmitted at only a half degree or so with reference to the surface of the earth. Therefore, the tropospheric wave may be classified as a space wave. Although duct transmission seems to be highly desirable, it is not frequently

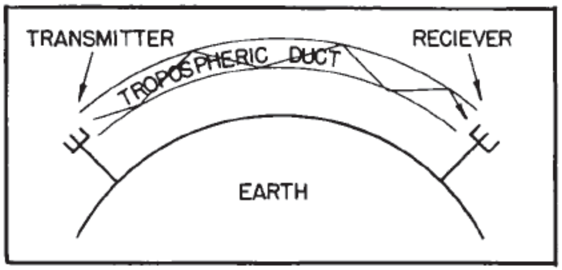


Figure 28-47 - The tropospheric duct.

used because of the unpredictable nature of the ducts. Temperature inversions occur frequently and at random times, with few exceptions. If the weather conditions could be accurately predicted over long distances, communications between two distant points utilizing the ducts would be possible with small amounts of power.

Not all frequencies can be transmitted through these ducts. The height of the duct above the earth determines the minimum frequency of transmission. If the height of the duct is only a few feet above the earth, the transmission by this means would be limited to the frequencies in the ultra high or super high portion of the frequency spectrum.

Q21. Is transmission strictly by use of the tropospheric duct a practical and reliable means of communications? Why?

Q22. Why is tropospheric transmission considered to be space wave transmission?

28-20. The Sky Wave

One of the most frequently used methods of long distance transmission is by the use of the SKY WAVE. Sky waves are those waves radiated from the transmitting antenna in a direction that produces a large angle with reference to the earth. The sky wave has the ability to strike the ionosphere, be refracted from it to the ground, strike the ground, be reflected back toward the ionosphere, and so forth. The refracting and reflecting action of the ionosphere and the ground is called SKIPPING. An illustration of this skipping effect is shown in Figure 28-48.

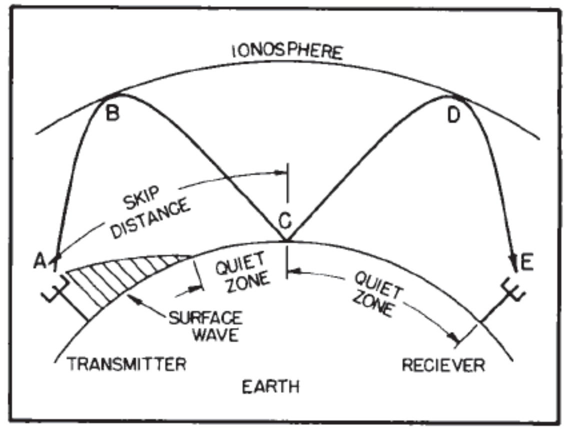


Figure 28-48 - Sky wave skipping and skip distance.

The transmitted wave leaves the antenna at point (A), is refracted from the ionosphere at point (B), is reflected from the ground at point (C), is again refracted from the ionosphere at point (D) and arrives at the receiving antenna (E). The points from (A) to (C), and from (C) to (E) indicate a distance that is known as the SKIP DISTANCE. The region from the end of the surface wave to point (C) is known as the SKIP or QUIET ZONE because a receiver located within these regions would receive none of the transmitted wave. It is possible, however, that a propagated wave leaving the antenna at a greater angle than the angle shown in Figure 28-48, would conceivably be refracted into that region. The chance of this happening is small because the angle at which the sky wave strikes the ionosphere is critical. The critical nature of the sky waves and the requirements for refraction will be discussed thoroughly later in this section. Notice that there is another quiet zone between points (C) and (E).

To understand the process of skipping, the composition of the atmosphere and the factors that affect it must be considered. As far as electromagnetic radiation is concerned, there are only three layers of the atmosphere. They are: the troposphere (already discussed), the STRATOSPHERE, and the ionosphere. The troposphere extends from the surface of the earth to an altitude of approximately 6. 5 miles. The next layer, the stratosphere, extends from the upper limit of the troposphere to an approximate elevation of 23 miles. From the upper limit of the stratosphere to a distance of approximately 250 miles lies the region known as the ionosphere. The temperature in the stratosphere is considered to be a constant unfluctuating value. Therefore, it is not subject to temperature inversions nor can it cause significant refractions. The constant temperature stratosphere is also called the ISOTHERMAL REGION.

The ionosphere is appropriately titled because it is composed primarily of ionized particles. The density at the upper extremities of the ionosphere is very low and becomes progressively higher as it extends downward toward the earth. The upper limit of the ionosphere is subjected to severe radiation from the sun. This radiation from the sun is in the form of

A18. Diffraction is the bending of a wave as it passes the edge of an obstruction. Refraction is the bending experienced by a wave as it moves between different density mediums.

A19. Optical horizon is the distance visible from the transmitting antenna to the actual horizon. Radio horizon is the maximum distance from the transmitting antenna to the receiving antenna over which reception can occur by direct waves. This is usually greater than the optical horizon due to the effects of refraction and diffraction.

A20. They are in phase.

A21. No. It is subject to fluctuations in weather conditions.

A22. Because of the comparatively low height of the tropospheric ducts.

---

photons, gamma rays, and other high energy particles. Even though the density of the gases in the upper ionosphere is small, the radiation particles from space are of such high energy that they cause wide scale ionization of the gas atoms that are present. This ionization extends down through the ionosphere with diminishing intensity. Therefore, the highest degree of ionization occurs at the upper extremities of the ionosphere, while the lowest degree occurs in the lower portion of the ionosphere.

The ionosphere is composed of three layers designated respectively from lowest level to highest level D, E, and F. The F layer is further divided into two layers designated F_1 (the lower layer) and F_2 (the higher layer). The presence or absence of these layers in the ionosphere and their height above the earth vary with the position of the sun. At high noon, radiation in the ionosphere directly above a given point is greatest, while at night it is minimum. When the radiation is removed, many of the particles which were ionized recombine. The interval of time between these conditions finds the position and number of the ionized layers within the ionosphere changing. Since the position of the sun varies with respect to a specified point on earth daily, monthly, and yearly, the exact position and number of layers present is extremely difficult to determine. However, the following general statements can be made:

The D layer ranges from about 25 to 55 miles. Ionization in the D layer is low because it is the lowest region of the ionosphere. This layer has the ability to refract signals of low frequencies. High frequencies pass right through it and are attenuated in so doing. After sunset the D layer dissappears because of the rapid recombination of ions.

The E layer limits are from approximately 55 to 90 miles high. This layer is also known as the Kennelly-Heaviside layer, because these two men were the first to propose its existance. The rate of ionic recombination in this layer is rather rapid after sunset, and is almost gone by midnight. This layer has the ability to refract signals of a higher frequency than was refracted by the D layer. In fact, the E layer can refract signals as high as 20 Mc. For this reason, it is valuable for communications in ranges up to about 1, 500 miles.

The F layer exists from about 90 to 240 miles. During the daylight hours, the F layer separates into two layers, the F_1 and F_2 layers. The ionization level in these layers is quite high and varies widely during the course of a day. At noon, this portion of the atmosphere is closest to the sun and the degree of ionization is maximum. Since the atmosphere is rarefied at these heights, the recombination of the ions occurs slowly after sunset. Therefore, a fairly constant ionized layer is present at all times. The F layers are responsible for high frequency long distance transmission.

The relative distribution of the ionospheric layers is shown in Figure 28-49. With the disappearance of the D and E layer at night, signals normally refracted by this layer are refracted by the much higher layer, resulting in greater skip distances at night, as shown in Figure 28-50.

The layers which form the ionosphere undergo considerable variations in attitude, density and thickness, due primarily to varying degrees of solar activity. The F_2 layer undergoes the greatest variation due to solar disturbances (sun-spot activity). There is a greater concentraction of solar radiation in the earth's atmosphere during peak sun-spot activity which recurs in 11 year cycles (minimum, 1955; maximum, 1959-1960).

During periods of maximum sun-spot activity, the F layer is more dense and occurs at a higher attitude, as shown in Figure 28-51. In Figure 28-51, conditions are shown for transmitted wavefronts A and B having different angles of radiation. During periods of minimum sun-spot activity, the lower altitude of the F layer returns the skywaves (dotted lines) to points relatively close to the transmitter compared with the higher altitude F layer occurring during maximum sun-spot activity. Consequently, skip distance is affected by the degree of solar disturbance.

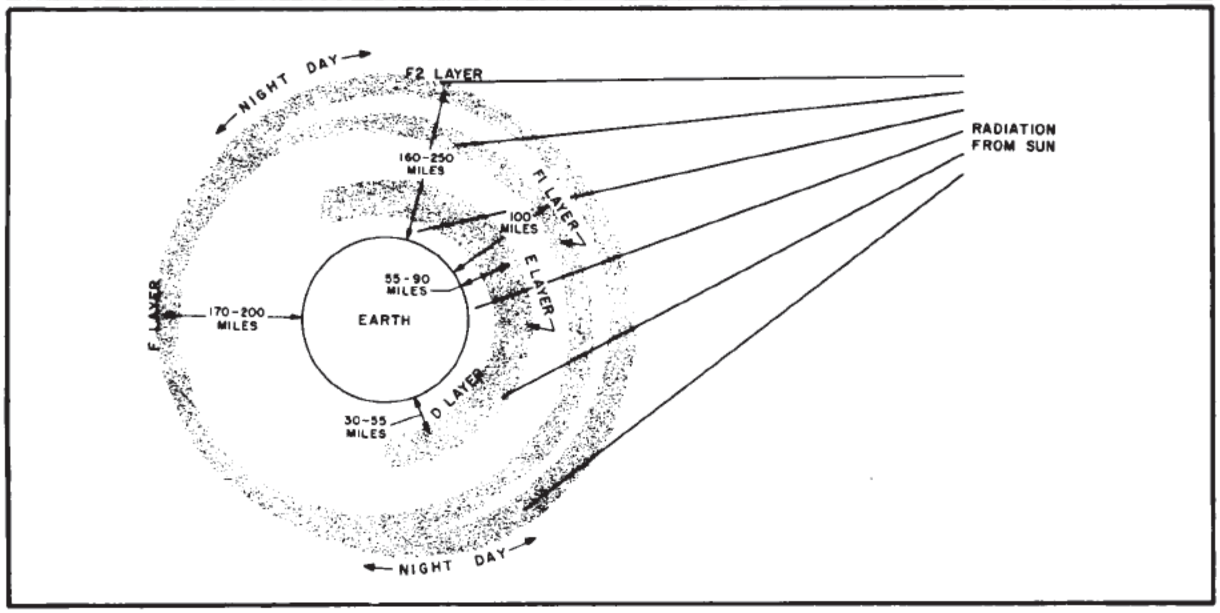


Figure 28-49 - Layers of the ionosphere.

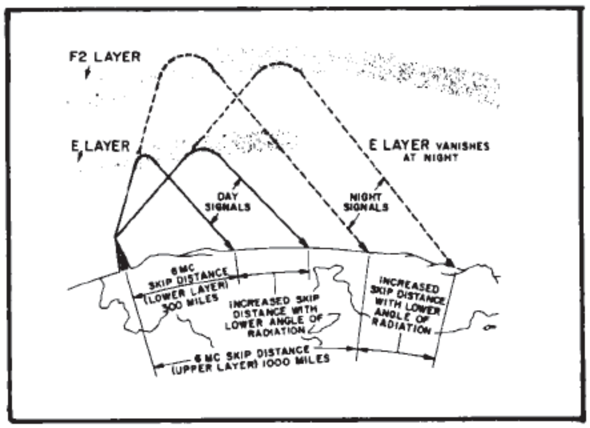


Figure 28-50 - Effect on skip distance with disappearance of E layer.

28-21. Effects of the Ionosphere on the Sky Wave

The ionosphere has many characteristics. Some waves penetrate and pass entirely through it into space, never to return. Other waves penetrate but bend. Generally, the ionosphere acts as a conductor, and absorbs energy in varying amounts from the radio wave. The ionosphere also acts as a radio mirror and refracts (bends) the sky wave back to the earth, as illustrated in Figure 28-52. Here, the ionosphere does by refraction similar to what water does to a beam of light.

The ability of the ionosphere to return a radio wave to the earth depends upon the ion density, frequency of transmission and the angle of radiation. The refractive power of the ionosphere increases with density or degree of ionization. The degree of ionization is greater

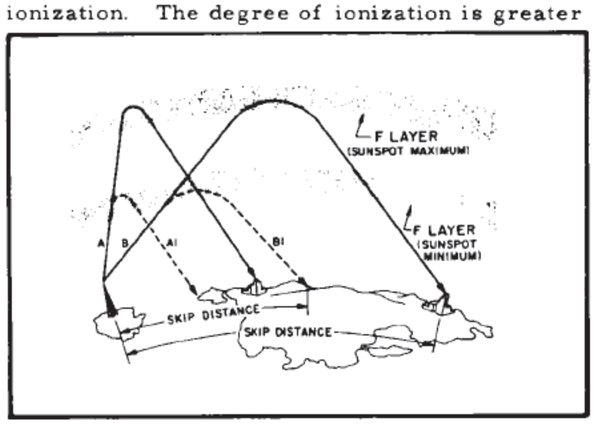


Figure 28-51 - Effects of sun-spot activity on skip distance.

in summer than in winter, and is also greater during the day than at night. As mentioned

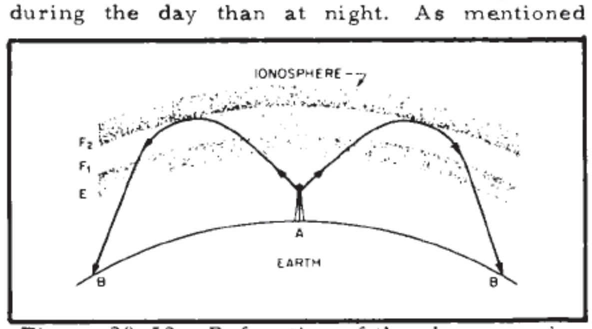


Figure 28-52 - Refraction of the sky waves by the ionosphere.

previously, abnormally high densities occur during times of peak sun-spot activity

If the frequency of a radio wave being transmitted vertically is gradually increased, a point will be reached where the wave will not be refracted sufficiently to curve their path back to earth. Instead, these waves continue upward to the next layer where refraction continues. If the frequency is sufficiently high, the wave will penetrate all layers of the ionosphere and continue on out into space. The highest frequency which will be returned to earth when transmitted vertically under given ionospheric conditions is called the CRITICAL FREQUENCY.

In general, the lower the frequency, the more easily the signal is refracted; conversely, the higher the frequency, the more difficult is the refracting or bending process. Figure 28-53 illustrates this point.

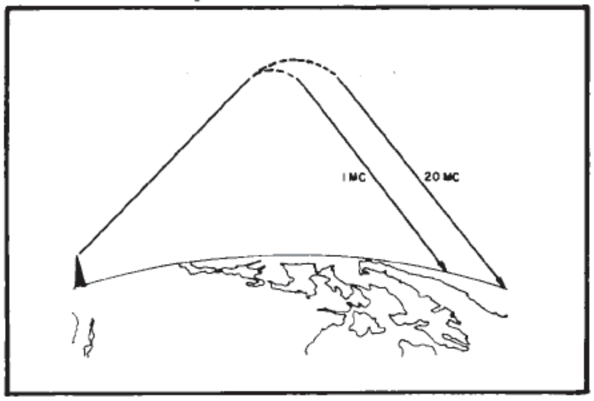


Figure 28-53 - Relationship of frequency to refraction by the ionosphere.

The angle of radiation plays an important part in determining whether a particular frequency will be returned to earth by refraction from the ionosphere. Above a certain frequency, waves transmitted vertically continue on into space. However, if the angle of propagation is lowered (made less vertical), a portion of the high frequency waves will be returned to earth. The highest angle at which a wave can be propagated and still be returned from the ionosphere is called the CRITICAL ANGLE for that particular frequency. For purposes of calculation, the critical angle is the angle which the wavefront path, at incidence with the ionosphere, makes with a line extended to the center of the earth.

Figure 28-54 is a light-beam analogy to demonstrate the critical angle concept. In the figure, a light source is shown well below the surface of the ocean. When the light source generates beam A, the light is refracted slightly and escapes the medium of water. With the light source tilted to the right, beam B is refracted a sufficient amount to cause the beam to skirt the surface of the water. Further tilting of the light source results in considerable refraction so that the beam returns to the ocean floor. In this analogy, angle θ_2 is the critical angle.

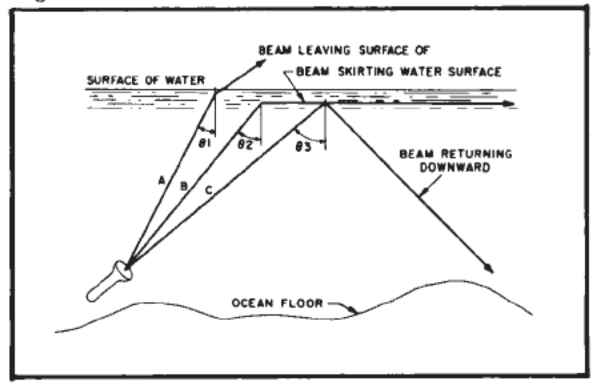


Figure 28-54 - Light-beam analogy, showing the effects of refraction and the critical angle.

The action of a radiated wave from an antenna is similar to that described for the light analogy above. When electromagnetic waves enter the ionosphere, they are effectively speeded up and follow curved paths, as shown in Figure 28-55. For the particular frequency used, angle θ represents the critical angle. Any wave that leaves the antenna at an angle greater than θ will penetrate the ionosphere and continue on into space. At angles less than the critical angle, the waves are refracted back to earth. It should be noted that the skip distance reduces as the angle of propagation increases toward the critical angle.

From the previous discussion it is obvious that there is a "best frequency." As can be seen in Figure 28-56, the distance between the transmitting antenna and the point at which the wave returns to earth depends upon the angle of propagation, which in turn is limited by the frequency. The highest frequency which is returned to earth at a given distance is called

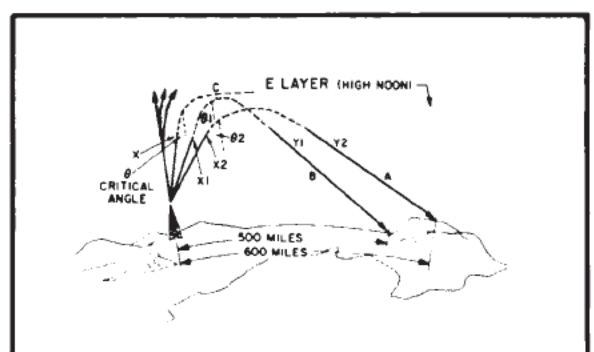


Figure 28-55 - Ionospheric refraction at various angles of propagation.

the MAXIMUM USABLE FREQUENCY (MUF) and has an average monthly value for any given time of the year. The OPTIMUM WORKING FREQUENCY is the one which provides the most consistent communication. For transmission using the F_2 layer, the optimum working frequency is about 85% of the MUF while transmission via the E layer will be consistent, in most cases, if a frequency near the MUF is used.

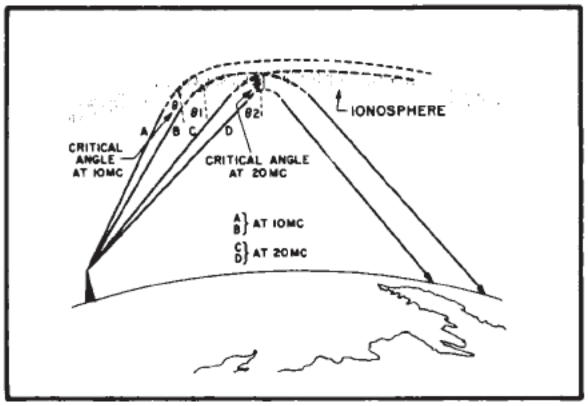


Figure 28-56 - Relationship of frequency to critical angle.

Because of this variation in the critical frequency, nomograms and frequency tables are issued that predict the maximum usable frequency (MUF) for every hour of the day for every locality in which transmission are made.

Nomograms and frequency tables are prepared from data obtained experimentally from stations scattered all over the world. All this information is pooled, and the results are tabulated in the form of long range predictions that remove most of the guess work from radio communications.

The increased ionization during the day is responsible for several important changes in sky wave transmission. It causes the sky wave to be returned to the earth nearer to the point of transmission. The extra ionization increases the absorption of energy from the sky wave. If the wave travels a sufficient distance into the ionosphere, it will lose all of its energy. The presence of the F_1 and E layers with the F_2 layer make long range high frequency communications possible, provided the correct frequencies are used.

Absorption usually reduces the effective daylight communication range of low frequency and medium frequency transmitter to surface wave ranges. The high degree of ionization of the F_2 layer during the day, enables refraction of high frequencies which are not greatly absorbed, has frequency band.

Due to the interplay of frequency and angle of propagation, long distance communication must take into account both factors. The following information indicates the approximate angle of radiation most suitable for radio waves of different frequencies and for different distances between points of communications.

1.5 to 3 Mc—low angle radiation for long distances. High angle radiation may cause fading of ground wave reception (to be discussed further, shortly). Vertical antennas are preferable.

3 to 7 Mc—Good sky wave return at any angle of radiation. High angle radiation can be used for short to moderate ranges, but low angle radiation should be used for long distance communications.

7 to 12 Mc—Angle of radiation from 45 to 30 degrees for short to moderate distances. Lower angles should be used for long distance communication. Higher radiation angles can be used to overcome variations in ion density during peaks of sun spot activity.

12 to 30 Mc—Not useful for short distance, sky wave transmission. The maximum useful angle when operating on a frequency of 12 to 16 Mc is about 30°. As the frequency is increased to 28 Mc, the angle of propagation should be decreased to 10 degrees. Above 28.5 Mc, an angle less than 10 degrees should be used.

It will be remembered that the ionosphere is composed of several layers, specifically D, E, F_1 and F_2, with the D and E layers becoming practically nonexistent at night while the F_1 and F_2 layers combine into a single layer at night. A given transmitted wave which uses the E layer for propagation will travel to the F layer at night, resulting in greater range of communication. The distance between the transmitting antenna and the point where a usable refracted wave is returned to the earth is called the skip distance. In a way of summary, then, it can be said that skip distance is affected by transmitted frequency, angle of propagation and changes in ionospheric conditions.

Between the point where the ground wave is completely dissipated and the point where the first sky wave returns NO signal will be heard. This area is called the quiet or skip zone and is shown in Figure 28-57.

A radio wave may be refracted many times between the transmitter and receiver locations. as shown in Figure 28-58. In this example the radio wave strikes the earth at location A, with sufficient intensity to be reflected back to the ionosphere and there to be refracted and returned to the earth a second time. Frequently a sky wave has sufficient energy to be refracted and reflected several times, greatly increasing the range of transmission. Because of this so-called multiple-hop transmission, transoceanic and around the world communication is possible with moderate power.

774-422 O - 65 - 9

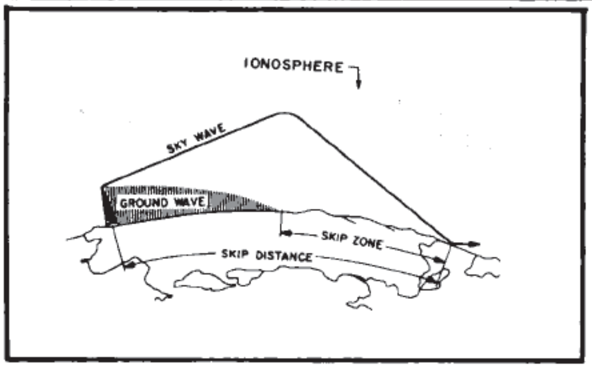


Figure 28-57 - Skip zone.

Fading is a term used to describe the variations in signal strength that occur at a receiver during the time a signal is being received. Fading may occur at any point where both the ground

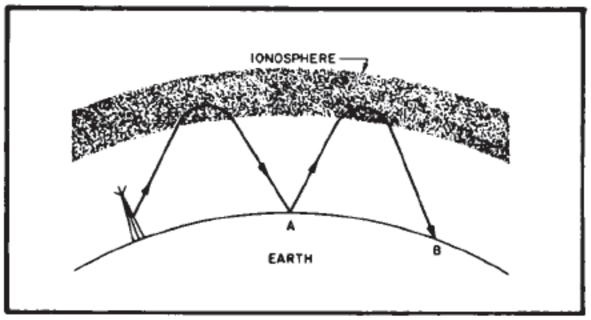


Figure 28-58 - Multiple refraction and reflection of a sky wave.

wave and the sky wave is received, as shown in Figure 28-59A. The two waves may arrive out of phase, thus producing a cancellation of the usable signal. This type of fading is encountered in long range communication over bodies of water where ground wave propagation extends for a relatively long distance.

In areas where sky wave propagation is prevalent, fading may be caused by two sky waves traveling different distances, thereby arriving at the same point out of phase, as shown in Figure 28-59B. Such a condition may be caused by a portion of the transmitted wave being refracted by the E layer while another portion of the wave is refracted by the F layer. A complete cancellation of the signal would occur if the two wave arrived 180° out of phase with equal amplitudes. Usually one signal is weaker than the other and, therefore, a usable signal is obtained.

Variations in absorption and in the length of the path in the ionosphere are also responsible for fading. Occasionally, sudden disturbances in the ionosphere cause complete absorption of all sky wave radiation.

Receivers located near the outer edge of the

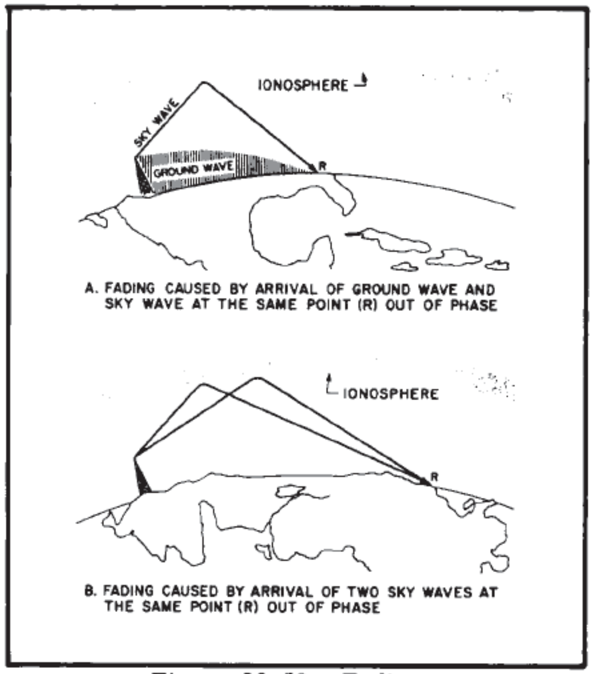


Figure 28-59 - Fading.

skip zone are subjected to fading as the sky wave alternately strikes and skips over the area. This type of fading sometimes causes the received signal strength to fall to nearly the zero level.

One method used to help reduce the effects of fading is to place two or more receiving antennas a wavelength or two apart, each antenna feeding its own receiver with all receiver audio outputs combined. This process is known as DIVERSITY RECEPTION. The greater the number of antennas and receivers used, the smaller the effects of fading becomes.

Frequency blackouts are closely related to certain types of fading, some of which are severe enough to completely blank out the transmission.

The changing conditions in the ionosphere shortly before sunrise and shortly after sunset may cause complete blackouts at certain frequencies. The higher frequency signals pass through the ionosphere while the lower frequency signals are absorbed by it.

Ionospheric storms (turbulent conditions in the ionosphere) often cause radio communication to become erratic. Some frequencies will be completely blacked out, while others may be reinforced. Sometimes these storms develop in a few minutes, and at other times they require as much as several hours to develop. A storm may last several days.

When frequency blackouts occur, the operator must be alert if he is not to lose contact with other ship or shore stations. In severe

storms the critical frequencies are much lower, and the absorption in the lower layers of the ionosphere is much greater.

Q23. What is the difference in height between the F_1 layer present during the day, and the F layer present at night?

Q24. What is the most obvious advantage to using sky wave transmission?

28-22. Navy Frequency Utilization

Each frequency band has its own special uses. These uses depend upon the nature of the waves (surface, sky, or space) and the effect that the sun, earth, and the atmosphere have on them.

It is difficult to establish fixed rules for the choice of a frequency for a particular purpose. Some general statements can be made, however, as to which frequency bands are best suited for the type of transmission to be made.

For example, if a long range communication is to get through to a distant receiver, high power and low frequency should be used. The large international communication systems and the fleet broadcast stations use this combination of low frequency and high power. (Fleet broadcase stations send their messages without requiring a reply by the stations receiving the messages). However, this combination requires an antenna array that may be too large for use with shipboard transmission. An alternative method is to transmit the message to the nearest shore station for relay to its destination.

During the day, the sky wave builds up to a peak of usefulness extending into the HF band, while at night the peak of the sky wave usefulness is reduced to the top third of the MF band. The usefulness of the ground or surface wave declines steadily as the higher frequencies are reached until it is of no value in the HF band. The only means of radio communication in the VLF band is the surface wave component of the ground wave.

Sky wave transmission (1, 600 to 30, 000 kc) is is almost always associated with skip distances. Great range can be obtained, but in the process many receiving stations may be skipped between the source and the most remote points of expected reception. Thus, one of the stations to which it is desired to get the signal may be skipped if the receiving station is located within any one of the skip zones associated with the frequency being transmitted.

The most important frequencies for long range transmission are those from 2, 000 to 18, 100 kc. This band is standard for long distanc- -aval communications from ship to ship ana from ship to shore. It is the band that is most frequently used and the one that is covered by the standard Navy transmitter such as the AN/SRT-14, 15 and 16, and AN/URC-32. The band is in the short wave region, and transmissions are accomplished by means of the sky wave and are affected by skip distances. When long range is desired during the day, the frequencies that should be used are from approximately 7 Mc to 18 Mc. For night communications, frequencies below 10 Mc to 15 Mc should be used.

When frequencies above the the band are transmitted, the use of sky waves is virtually eliminated. At certain times of the day all ionospheric layers exist; however, even under these conditions, waves above 30 Mc seldom receive sufficient refraction to return them to the earth.

The use of a surface wave for frequencies above the HF band results in a large amount of energy being dissipated in the form of heat, due to eddy currents induced in the ground as the wave travels along the earth's surface. The energy dissipated within the ground reduces the power content of the wave. At low frequencies, the power loss is quite small, but for frequencies above the HF band, the loss becomes excessive. As a result, surface waves having frequencies above the HF band become greatly attenuated at small distances from the transmitter.

Due to the effects just mentioned, propagation of waves at frequencies above the HF band is limited to the use of direct or space waves, frequently called line-of-sight transmission.

In recent years there has been a trend toward the use of frequencies above 30 Mc for short range ship to ship and ship to aircraft communications.

Early concepts suggested that these transmissions traveled in straight lines. This leads to the assumption that the VHF and UHF transmitters and receivers must be within sight of each other to maintain radio contact.

Extensive use of these frequencies coupled with additional research show that the early line-of-sight theory is frequently in error because radio waves of these frequencies may be refracted. The receiver does not always have to be in sight of the transmitter. Although this type of transmission still is called line-of-sight transmission, it is more meaningful to call it VHF and UHF transmission.

In general, the VHF and UHF waves follow approximately straight lines with large hills or mountains casting a radio shadow over areas in much the same way that they cast a shadow in the presence of light rays. A receiver located in a radio shadow will receive a weakened signal, and, in some cases, no signal at all. Theoretically, the range of contact is the distance to

A23. The F layer apparently rises at night because of the recombination, so that it is physically higher than the F_1 layer.

A24. Long distance transmission..

---

the horizon, which is determined by the heights of the two antennas. However, as stated previously, communication is sometimes possible many hundreds of miles beyond the assumed horizon range. This fact must be observed when transmission is of a security nature.

Nonregistered Publications Memoranda (DNC-14A), which are supplied to the various ships of the Navy, contain the tables that show the best frequencies within the band for communications with various shore stations. These tables give the recommended frequency for every hour of the day covering distances from 250 to 5,000 miles for some stations. The direction of the receiving station from the ship transmitting the signals is also taken into account. DNC-14A covers a three month period with a separate table for each month and for each major shore station. These tables are consulted for specific and current information.

Q25. What is the likely method of propagation used with a transmitted wave of 100 Mc?

EXERCISE 28

1. Describe the purpose of an antenna.
2. What is a dipole?
3. Describe the difference between the Hertz and Marconi antennas.
4. What are the radiation characteristics of a quarter-wave section of transmission line?
5. What is the impedance value at the receiving and sending end of an open quarter-wave transmission line?
6. Describe the distribution of current, voltage, magnetic and electrostatic lines of force on a quarter-wave section of transmission line.
7. How is a half-wave antenna constructed from a quarter-wave section?
8. What is a feed line?
9. What is the difference between center-fed and end-fed?
10. What is the difference between current-fed and voltage-fed?
11. How may small irregularities in the length of an antenna be corrected?
12. What type of transmission line is usually connected to an antenna?
13. What are the disadvantages in connecting a transmission line to an antenna?
14. What is a delta match?
15. What is the advantage of a delta match?
16. State the five different types of nonresonant lines that are used to feed a half-wave antenna.
17. What are the characteristics of a nonresonant line used to feed an antenna?
18. How does the electrical length of any given dipole compare to its physical length?
19. What is the physical length of an antenna built for 13 Mc?
20. Describe how an antenna may act if it is either too long or too short.
21. How can discrepancies in line length be compensated for?
22. Describe in your own words how propagation is caused.
23. What is the relationship of the induction field to the radiation field?
24. What is the relationship of the E and H fields in the induction field?
25. What is the relationship between the E and H fields in the radiation field?
26. What is the speed of propagated radio frequency energy?
27. What is meant by the term horizontal polarization?
28. What is meant by the term vertical polarization?
29. Give examples of vertical and horizontal polarization.
30. What is elliptical polarization?
31. Describe the non-uniformity of the fields that exist about an antenna.
32. What is a lobe?
33. What is a null?
34. What is meant by the term directivity?
35. How is antenna input impedance computed?
36. What is radiation resistance?
37. Compare the input impedance of a half-wave thin and thick diameter antenna.
38. What is the Marconi antenna?
39. How is the Marconi antenna polarized?
40. What is meant by counterpoise?
41. What is the purpose of a counterpose?
42. Describe the advantages of the L antenna as compared to a vertical antenna of similar height.
43. What is meant by the term antenna gain?
44. What is an isotropic antenna?
45. How may the gain in db's of an antenna be expressed if its power ratio is known?
46. What is a long wire antenna?
47. Compare the long wire antenna to the half-wave antenna.
48. What is the value of a radiation pattern?
49. Why is an antenna tuning circuit necessary?
50. What is meant by the term free space?
51. What is the radio frequency spectrum?
52. What is meant by the abbreviation MF?
53. What are the two most general classifications of propagated waves?
54. What is a surface wave?
55. How is a surface wave affected by contact with soil of extremely low conductivity?
56. What is meant by the term angle of tilt?
57. What is the surface wave commonly used for?
58. What is a direct wave?
59. What is a space wave?
60. What is the relationship between the ground reflected wave and the direct wave?
61. What is meant by the term line-of-sight?
62. What is refraction?
63. What is diffraction?
64. Compare the optical horizon with the radio horizon.
65. What is the relationship between a ground incident wave and a ground reflected wave?
66. What is the relationship between the direct wave and the ground reflected wave at the receiver antenna?
67. How does reflection affect polarization?
68. What is a shadow zone?
69. What is a tropospheric duct? How is it formed?

A25. Space wave propagation since, at this frequency, surface waves would produce excessive power losses while sky waves would penetrate the ionosphere and become lost.

70. What is a sky wave?
71. How is the sky wave of medium frequency affected by the ionosphere?
72. How is the ionosphere created?
73. Describe all the layers of the ionosphere in terms of their height, density, and presence during various periods of one day.
74. What is meant by the term skip distance?
75. What is a skip zone?
76. Describe how different frequencies are affected by the different layers of the ionosphere?
77. What is meant by the terms critical frequency and critical angle?
78. What are the most important frequencies used by the Navy in long range transmission?

D

E

F

U.S. GOVERNMENT PRINTING OFFICE : 1965 O - 774-422